



HAL
open science

L'état du droit : expérience de reconnaissance automatique des jurisprudences de la justice administrative

Alioscha Massein

► **To cite this version:**

Alioscha Massein. L'état du droit : expérience de reconnaissance automatique des jurisprudences de la justice administrative. Informatique [cs]. 2022. dumas-03790883

HAL Id: dumas-03790883

<https://dumas.ccsd.cnrs.fr/dumas-03790883v1>

Submitted on 28 Sep 2022

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution 4.0 International License

L'état du Droit

Expérience de reconnaissance
automatique des jurisprudences de la
justice administrative

Sous la direction de Sabine Loudcher

Remerciements	2
Introduction	3
I. De la décision de justice aux données juridiques : constituer le corpus	12
A. Historique et questionnements autour de la numérisation des documents juridiques	13
B. Les décisions de justice : de la décision aux données.	16
1. <i>La base de données JADE et les décisions de justice administrative</i>	16
2. <i>Les métadonnées</i>	19
3. <i>Le texte de la décision</i>	21
C. Comment utiliser ces données ? Sélectionner l'information	24
1. <i>Sur les métadonnées</i>	24
2. <i>Sur le texte des décisions</i>	27
II. Des données juridiques à la science du Droit : l'apport du numérique et le poids des choix	30
A. Traiter le corpus pour le rendre analysable : vectoriser le texte	31
1. <i>Rendre le texte comparable</i>	31
2. <i>Automatiser la recherche des jurisprudences</i>	33
B. Exemples d'analyses	36
1. <i>Étude sur la concentration des considérants sur quelques jurisprudences.</i>	37
2. <i>Étude d'un cas particulier de jurisprudence : limites du travail d'analyse</i>	43
C. Une révolution (épistémologique) du Droit ?	48
1. <i>De « nouvelles » connaissances ?</i>	49
2. <i>Faire évoluer l'étude des jurisprudences et des contentieux</i>	52
3. <i>Enjeux du Droit par le prisme des outils numériques</i>	55
Conclusion	59
Bibliographie	62
Table des Matières	65
Annexes	67

Remerciements

Écrire ces lignes et plus largement d'avoir poursuivi mes études sur ces deux dernières années en Humanité Numérique n'ont pas été des choix simples sur des niveaux personnels et professionnels. Entre les doutes et les envies de tout lâcher, de manque de confiance dans ce parcours et face à la précarité des situations, je souhaiterais remercier ici toutes celles et ceux qui m'ont soutenu, compris, aidé ou qui ont simplement pris le temps de discuter de ces difficultés.

En premier, mes remerciements, de tout mon cœur, vont à Léna, qui m'accompagne chaque jour, et qui continue, avec toutes ces années, de m'apporter sa joie de vivre, ses rires et sa force, ou encore sa patience. Pour tout ça, et bien d'autres choses, je te remercie de m'avoir accompagné sur cette période qui n'aurait pas été possible sans toi. Je remercie aussi tous mes amis, proches, et familles qui m'ont également soutenu dans cette poursuite d'étude, et je pense en particulier à Benoît, dont la parole sincère a toujours su reconforter, et me permettre de remettre du contexte dans « les choses de la vie », à Cassandra, dont les apparitions occasionnelles sont toujours un vent de fraîcheur dans les canicules de la vie quotidienne, à Léna D, dont la ferveur n'a d'égale que les douces soirées au bar, à Clara et Andres, qui ont l'amour contagieux, à mes parents, et à mes frères, si différents et si précieux.

A l'instar de mes connaissances personnels, je tiendrais à remercier toutes les équipes universitaires, les professeurs, les personnes que j'ai pu rencontré dans les expériences professionnelles que j'ai pu réalisé sur ces deux dernières années. Je pense d'abord à Géraldine, dont les fou rires, les questionnements, et les longues discussions sur tout les sujets vont beaucoup me manquer. Je remercie également dans l'ensemble les personnes du SCUJO, qui travaille tous les jours, comme une grande partie du personnel de l'université, à faire en sorte de tenir, souvent à bout de bras, le fonctionnement de ce qu'il reste encore de cette institution publique nécessaire.

Je remercie également en dernier lieu les personnes sans qui ce stage et ce mémoire n'aurait pas été possible : à Isabelle Sayn, dont l'oreille attentive a su me diriger, et m'accompagner dans mes questionnements, à Marc Clément, dont la curiosité et l'écoute m'ont toujours surpris et amusé, à Julien, qui m'a beaucoup apporté malgré les peu de temps où nous avons pu nous voir. Enfin, à Sabine Loudcher, qui a pris le temps de m'écouter et de suivre ce mémoire sur toute la durée de sa production.

Pour toutes ces choses, et d'autres que j'oublie, je vous remercie toutes et tous.

Introduction

Au début de l'année 2022, un panel de journaux, revues et médias se sont questionnés sur la création en Chine d'une « intelligence artificielle » capable de prendre des décisions de justice sur un ensemble réduit de crimes et délits, et de décider si un individu est coupable des faits qui lui sont reprochés¹. Outre les nombreuses questions éthiques et techniques que posent ce programme et son utilisation, cette nouvelle application d'un modèle d'intelligence artificielle montre l'évolution que subit l'ensemble de la société au travers du numérique, et en particulier la Justice. Comme le dit Bachimont : « *Pour une part, cette situation est la suite logique du processus général de grammatisation auquel on assiste depuis les origines de l'Homme, où l'expression de son rapport au monde se manifeste à travers des supports matériels qui permettent d'objectiver sa représentation du monde* »². Comme le précise l'article, « *la Chine n'est pas le seul pays à s'intéresser à l'IA juridique et judiciaire* », la France et les États-Unis sont également deux pays qui cherchent à construire des outils informatiques qui ont des applications directes dans le monde juridique. Le fonctionnement du Droit, et par son application la Justice et les magistrat.es, est amené à être transformé par ces évolutions numériques.

« *Si le Droit saisit l'intelligence artificielle, afin d'encadrer son progrès et limiter ses dérives, il est encore plus vrai que l'intelligence artificielle saisit le Droit. Elle le remodèle, le redessine, mais ce nouveau monde juridique qui apparaît présente des contours encore flous. L'intelligence artificielle change à la fois les métiers et les marchés du Droit, les méthodes des juristes et les normes qu'ils appliquent ou étudient. L'influence des algorithmes sur le Droit est telle que la plupart des juristes sont (ou seront bientôt) touchés et donc concernés par les nouvelles « lois » produites par les traitements automatisés de données.* »³. Les outils informatiques ne sont pas seulement définis par le Droit, il existe également un ensemble d'acteurs (entreprises, magistrat.es, chercheur.ses, juristes,...)

¹ Slama, Rémi. « La Chine annonce avoir développé une intelligence artificielle capable de remplacer les juges ». *Laboratoire de cyberjustice*, <https://www.cyberjustice.ca/2022/02/01/la-chine-annonce-avoir-developpe-une-intelligence-artificielle-capable-de-remplacer-les-juges/>.

² Bachimont, Bruno. *Le nominalisme et la culture : questions posées par les enjeux du numérique*. p. 8.

³ Barraud, Boris. « Le Droit en datas : comment l'intelligence artificielle redessine le monde juridique (première partie : La dictature des algorithmes ou l'intelligence artificielle à la source du Droit) ». *Revue Lamy Droit de l'immatériel*, 2019. HAL Archives Ouvertes, <https://hal.archives-ouvertes.fr/hal-02073660>.

qui se saisissent de ces mêmes méthodes et outils pour créer des analyses sur le Droit ou des outils d'aide à la décision, qui transforment les « *marchés du Droit* ». Par cette citation, Boris Barraud nous amène à réfléchir à l'importance qu'ont les données, et les techniques statistiques dans l'univers du Droit et de la Justice. En particulier dans l'espace qui nous concerne : la science du Droit⁴, c'est-à-dire pour nous le « *Droit saisi par la technologie* »⁵, sur laquelle nous reviendrons. Ce constat est également partagé par tout un ensemble d'auteurs et d'autrices, en particulier Danièle Bourcier, qui a pointé très tôt les innovations que peut apporter l'informatique au Droit⁶, mais également Pascale Deumier, qui revient plus largement sur les évolutions qui peuvent amener le travail sur la « masse » des « décisions du fond ». Serge Bories a également contribué à promouvoir cette approche des techniques informatiques sur des matériaux juridiques. Ces questionnements et ces possibilités peuvent même remonter jusqu'à la fin des années cinquante, notamment dans l'article de Jean Goulet, *Revalorisation du Droit et jurimétrie*, où l'auteur s'applique à montrer dans quelle mesure l'informatisation des contenus permet à la fois un gain de coût en terme de stockage de l'information, mais également dans les potentiels que permettraient la numérisation des documents juridiques : en termes de coût pour les tribunaux, mais également en termes d'aide à la décision pour les avocats et la justice. « *L'ordinateur électronique coûte cher. C'est vrai ! Est-il besoin pourtant, à partir des statistiques que nous venons d'élaborer, de recourir à une évaluation précise du prix de revient de l'entreposage et de l'utilisation de la masse de documents légaux que nous connaissons, pour démontrer que la méthode traditionnelle de recherches juridiques représente une dépense encore plus considérable ? [...] L'ordinateur ne rendra jamais la justice, mais il aide à la rendre. S'il accomplit ce rôle, il a aussi par conséquent sa place au sein de la salle d'audience* »⁷.

Plusieurs points sont ici à détailler : d'abord, il s'agit de bien préciser que notre démarche est de traiter de « *science du Droit* », au sens que lui donne Danièle Bourcier, et non de l'informatique décisionnelle en Droit, en particulier au travers de méthodes d'apprentissage automatique (*machine learning*). Nous verrons que cet espace de la

⁴ Bourcier, Danièle. *À propos des fondements épistémologiques d'une science du Droit*. Presses Universitaires de France, 2007. www-cairn-info.bibelec.univ-lyon2.fr, <http://www.cairn.info/quelles-perspectives-pour-la-recherche-juridique--9782130558989-page-69.htm>.

⁵ Barraud, Boris, *Art. Cit.*

⁶ Bourcier, Danièle, *Ibid.*

⁷ Goulet, Jean. « Revalorisation du Droit et jurimétrie ». *Les Cahiers de Droit*, vol. 9, n° 1, 1967, p. 9-36. www.erudit.org, <https://doi.org/10.7202/1004340ar>.

recherche se rapproche des problématiques des humanités numériques. Ensuite, les auteurs et les autrices précédent.es abordent d'autres sujets qu'il nous faudra parcourir, comme la question des « marchés du Droit », le traitement des données « massives », ou encore les questions que nous nous sommes posées à partir de ce « *gisement d'informations judiciaires* »⁸.

Qu'est-ce que la science du Droit par rapport à la « doctrine », traditionnellement étudiée au sein des universités ? Pour Danièle Bourcier : « *Le Droit est un ensemble de règles et de décisions dans un ordre juridique donné. La science du Droit prend le Droit comme objet de recherche. Cet objet est construit à l'aide du modèle qui le décrit. La science du Droit doit être distinguée de la Doctrine juridique, qui fait appel à l'opinion et aux commentaires des jurisconsultes sur le Droit positif* »⁹. La science du Droit, en tant que discipline ne se constitue qu'à la fin des années quatre-vingt-dix, et c'est notamment avec l'essor de l'informatique que se légitime cet espace de la recherche : « *Les auteurs, Buchanan et Headrick, après avoir reproché au juriste de ne voir dans les ordinateurs que des moyens de stocker leurs textes, donnent une autre vision de ce que pourrait être un véritable travail scientifique interdisciplinaire sur le Droit : tenter « d'incorporer (incorporate) dans l'ordinateur les techniques complexes du raisonnement juridique ».* C'est bien ce travail qui s'est développé dans notre discipline depuis plus de trente ans. ». Dans une perspective épistémologique, l'autrice tente de montrer que les sujets du Droit leurs sont propres, mais que les outils utilisés, qu'ils soient linguistiques ou informatiques, sont indépendants du champ de la recherche en science du Droit. C'est en ce sens que la science du Droit est nécessairement pluridisciplinaire, et, de ce point de vue, rejoint les pratiques des humanités numériques. L'utilisation croissante de l'informatique dans notre société, et *a fortiori* dans le monde du Droit et de la justice invite « *les juristes à la réflexion face au développement des technologies nouvelles, aux nouveaux supports d'informations, au processus de traitement de l'information et à ses apports prévisibles à la connaissance du Droit et à la recherche juridique, enfin, aux nouveaux comportements qu'ils induisent. Car le support n'est pas neutre et notre raison semble bien dépendre en partie de nos outils* »¹⁰.

⁸ Bories, Serge. « L'informatisation des données judiciaires et doctrinales : une contribution à la connaissance et à la recherche juridiques ». *Documentaliste-Sciences de l'Information*, vol. 40, n° 4, 2003, p. 272-79.

⁹ Bourcier Danièle, *Art. Cit*

¹⁰ Bories, Serge, *Ibid.*

Les humanités numériques peuvent jouer un rôle en science du Droit notamment car comme l'affirme Boulesnane et Bouzidi : « *elles s'appuient [...] sur l'ensemble des paradigmes, savoir-faire et connaissances propres à ces disciplines, tout en mobilisant les outils et les perspectives singulières du champ du numérique.(...) Les digital humanities désignent une transdiscipline, porteuse des méthodes, des dispositifs et des perspectives heuristiques liés au numérique* »¹¹. Les chercheur.ses et spécialistes en Humanités Numériques sont donc tout particulièrement indiqués pour les travaux qui reposent sur la manipulation des données en sciences humaines et sociales. D'abord parce qu'ils et elles sont à même de mettre en oeuvre un ensemble de techniques et d'outils informatiques adapté aux besoins du projet de recherche. Ils sont capables de constituer un corpus en portant une réflexion sur les échantillons de données pertinents pour les questions de recherche ainsi que sur la manière de « transformer » des données « brutes » en des données exploitables numériquement. Enfin, ils et elles sont capables de poser un regard critique sur les fonctions développées et implémentées dans le processus d'analyse de données, en ce qu'elles influencent nécessairement les résultats et les informations que l'on peut « extraire » des données : « *L'acteur humain intervient au niveau technologique, par le biais de l'usage et de la « modération » des TIC, et au niveau cognitif puisqu'il se base sur son référentiel, son expérience personnelle et professionnelle, son intuition, son raisonnement et sa logique d'interprétation pour la gestion des analyses* »¹². Enfin, les humanités numériques invitent tout particulièrement à éviter une réflexion trop positiviste liées à l'interprétation des données : entre « *un positivisme naïf où le fait collecté est la seule positivité à interroger et d'où il faut partir pour appliquer les modèles de construction scientifique* »¹³ et le rejet des données sous toutes leur formes du fait qu'elles essentialisent nécessairement des faits ou phénomènes sociaux et des individus. C'est dans cette tension permanente qu'une approche « humaniste » est particulièrement adaptée : il s'agit alors de prendre en compte les informations sous formes de données dans leur construction, leur transformation, leur exploitation, et les évolutions que ces opérations suggèrent sur les résultats, et les regards que l'on portent sur les données d'origines. Les chercheurs et chercheuses en humanité numérique sont ainsi des traducteur.trices, mais aussi des intermédiaires entre une approche savante d'un savoir spécifique - ici le Droit - et les outils informatiques :

¹¹ Bouzidi, Laïd, et Sabrina Boulesnane. « Les humanités numériques ». *Les Cahiers du numérique*, vol. 13, n° 3, 2017, p. 19-38.

¹² *Ibid.*

¹³ Bachimont, Bruno, *Art. Cit.*

« Cette mutation de statut consiste à transformer le contenu d'expression d'une pensée ou d'une intention en un fait constaté qui se traite désormais à travers des outils de collation, comparaison, transformation »¹⁴. C'est donc dans la pluridisciplinarité, dans la numérisation croissante des moyens et des matériaux, et les questionnements sur les données juridiques et leurs analyses que l'on retrouve de forte proximité entre les humanités numériques et les sciences du Droit.

La première condition de cette « hybridation » entre Droit et informatique, est la transformation de matériaux juridiques en matériaux informatiques, c'est-à-dire en données qui soient exploitables pour l'analyse statistique. C'est précisément ce que va permettre la loi du 7 octobre 2016 pour une République numérique, qui étend la mise à disposition publique de l'ensemble des jugements, par rapport à la loi du 17 juillet 1978 relative à l'accès aux documents administratifs. Si cette dernière permettait « d'obtenir communication des documents détenus par une administration, quels que soient leur forme ou leur support »¹⁵, elle ne mettrait pas en oeuvre le principe de publication des documents, qui permet leur ré-utilisation par d'autres acteurs. C'est ce qu'introduit la loi pour une République numérique, qui instaure l'ouverture par défaut des données administratives. Il en va de même avec les décisions de justice, qui sont publiées après anonymisation, et sont accessibles en ligne, notamment sur le site de la DILA, la Direction de l'Information Légale et Administrative. L'accès ainsi à une base de plus de 500.000 décisions de justice, des cours d'appels administratives et du conseil d'Etat, est un vaste ensemble de documents juridiques, sur lequel notre travail de stage portera. C'est aussi parce que cet ensemble de documents est quantitativement important que le travail statistique peut fonctionner, ou tout du moins peut être tenté. D'autant plus qu'à l'inverse des études classiques sur le Droit, le travail statistique sur la masse est récent, mais n'est pas forcément mis en valeur parmi les juristes : « On comprend alors que la doctrine distingue plusieurs sens au mot « jurisprudence », soulignant ainsi l'ambiguïté dont il souffre de ce fait : on se réfère principalement, d'une part, à la masse indistincte des innombrables décisions, contentieux présenté souvent comme sans intérêt pour la doctrine et d'ailleurs sans portée doctrinale, et, d'autre part, à ce qui serait en quelque sorte la « vraie » jurisprudence, issue de fort peu d'arrêts mais digne d'examen, d'exégèse, de diffusion et de mémorisation. Il y a, d'une façon sous-jacente, l'idée d'une

¹⁴ Bachimont, Bruno, *Art. Cit.*

¹⁵ « Loi pour une République numérique ». *Wikipédia*, 3 mai 2022. *Wikipedia*, <https://fr.wikipedia.org/w/index.php?title=Loi pour une R%C3%A9publique num%C3%A9rique&oldid=193367622>

jurisprudence plus noble que d'autre. »¹⁶. On retrouve cette idée d'une dichotomie dans la manière d'envisager le Droit chez Pascale Deumier¹⁷. Pour ces auteurs, il y a un intérêt scientifique à étudier cette masse : d'abord pour mesurer l'importance ou non de l'autonomie des juges de premières instances (« du fond »), par rapport au conseil d'Etat. Contrairement à l'idée selon laquelle le conseil d'Etat contrôle entièrement les juges des juridictions du fond et d'appels, les juges ont une certaine autonomie dans les décisions qu'ils et elles prennent. Serge Bories et Marie-Anne Frison-roche élaborent sur cette idée de manière résolument critique : « *Cette autonomie des décisions des juges du fond, de leur sens, de leur motivation et de leur résultat, est plus nette encore lorsque l'application jurisprudentielle d'une loi nouvelle n'a eu le temps de s'opérer qu'à l'échelon du premier ou second degré. Par la suite, lorsque la jurisprudence tutélaire de la Cour de cassation s'est exprimée, avec la part d'inflexibilité et d'appauvrissement corrélatif que cela implique nécessairement, disparaît cette richesse première. Plus une jurisprudence est récente, - le temps matériel de son contrôle, de son harmonisation, voire de sa normalisation n'étant pas encore déroulé -, plus elle est libre et créatrice de Droit.* »¹⁸. L'analyse de ces données peut également amener des intérêts pratiques, notamment pour le législateur, afin d'avoir une meilleure connaissance de l'application et des effets des lois qu'il promulgue. C'est donc l'analyse des décisions de justice de manière statistique qui peut permettre de mettre à jour des processus institutionnels, et de donner à voir des réalités sociales qui ne sont pas objectivement observées, sinon soupçonnées, par les acteurs et actrices de cet espace juridique. La numérisation des documents juridiques amènent cependant avec elle un ensemble d'inquiétudes et d'enjeux spécifiques : « *Les enjeux de la justice prédictive sont désormais bien connus. D'un côté, elle fait miroiter le développement d'un nouveau marché numérique, ce qui a motivé les amendements ayant introduit l'open data des décisions ; la transparence de la justice à l'égard du citoyen ; l'harmonisation de l'application du Droit sur le territoire ; le désencombrement des juridictions par l'augmentation des résolutions amiables ou le renoncement à une action vouée à l'échec. De l'autre, elle fait redouter la pression exercée sur la prise de décision des juges ; l'oubli de la spécificité des cas ; la possibilité de manipuler les algorithmes et les statistiques ; la transformation d'une part de la justice en marché* »¹⁹.

¹⁶ Frison-Roche, Marie-Anne, La jurisprudence massive, en collaboration avec Serge BORIES, D.1993, chron., p.287

¹⁷ Deumier, Pascale. « La justice prédictive et les sources du Droit : la jurisprudence du fond ». *Archives de philosophie du Droit*, vol. 60, n° 1, 2018, p. 49-66.

¹⁸ Frison-Roche, Marie-Anne, *Ibid.*

¹⁹ Deumier, Pascale, *Ibid.*

C'est dans ce cadre à la fois pluridisciplinaire et permis par l'*open-data* (ou données ouvertes) des décisions de justice mis en oeuvre à partir de 2016, que se transforment les « *marchés du Droit* », comme l'on a pu l'évoquer plus tôt avec le texte de Boris Barraud. Nous pouvons ici distinguer deux principaux « types » d'acteurs : les *legal-techs* - c'est-à-dire des entreprises privées qui développent des outils en direction des cabinets d'avocats - et les initiatives de recherches, qui sont, elles, publiques et financées, ou non, par des subventions publiques. Pour les deux catégories d'acteurs, il existe des motivations différentes à investir dans ces projets : pour les legal-tech, il s'agit probablement avant tout de chercher une rentabilité au travers du développement d'outils d'aide à la décision. Pascale Deumier relève d'ailleurs que celles-ci s'intéressent à des cas particuliers de documents juridiques : « *les outils actuellement développés, loin de traiter de la justice en général, ont en commun de s'intéresser à un niveau de décisions, celles du fond, et à un type de contentieux, celui qui est chiffré : indemnités de licenciement, pensions alimentaires ou dommages corporels sont les entrées unanimement explorées par les legal techs. Elles correspondent à des données plus faciles à traiter, à un contentieux répétitif – et à un marché à proportion.* »²⁰ Cependant, ces dernières ne s'intéressent que très peu à l'interprétation du Droit, à la dimension cognitive de celui-ci : « *le développement actuel d'outils de traitement du contentieux du fond ne s'intéresse qu'aux données chiffrées et aux contentieux portant sur ces données, comme les contentieux indemnitaires. La dimension cognitive est plus ténue et, selon plusieurs auteurs, bien plus difficile à appréhender.* ». L'auteur avance ainsi deux hypothèses qui peuvent expliquer ce phénomène : soit un nombre insuffisant de données, soit un marché trop peu intéressant. Nous avancerons ici une troisième hypothèse, qui nous paraît tout aussi importante : le désintérêt des legal-tech pour ce type d'étude serait en parti lié à l'absence de données **annotées** massivement, ce qui complexifie fortement la création de modèles d'intelligences artificielles, et peut être très coûteux. C'est aussi dans cette perspective que l'Etat, et le ministère de la justice ont lancé plusieurs projets pour annoter des décisions, comme le projet DataJust, récemment abandonné par le ministère de la Justice²¹.

De l'autre côté, les acteurs publics ont également des intérêts à financer des projets de recherche sur les données de la justice. Comme pouvait l'évoquer déjà Jean Goulet, l'informatique peut réduire le temps de la procédure, le travail de recherche et de

²⁰ Deumier, Pascale, *Art. Cit.*

²¹ Rivollier, Vincent, et al. « Le retrait de DataJust, ou la fausse défaite des barèmes ». *Recueil Dalloz*, n° 09, mars 2022, p. 467.

compilation des informations sur les précédents juridiques, ou encore le prix de stockage des informations. L'initiative DataJust montre aussi une volonté de l'Etat de développer et de soutenir des marchés privés. La question de l'efficience de la Justice et des démarches juridiques est aussi un enjeu pour les pouvoirs publics : il s'agit de savoir si certaines procédures peuvent éviter des passages au tribunal, et être en partie automatisées. Les magistrat.es eux-mêmes sont confrontés aux outils, et se questionnent sur l'évolution de leurs pratiques. En s'appuyant sur le texte de Marc Clément, magistrat au tribunal administratif de Lyon, celui-ci nous suggère non pas de regarder ce que pourrait transformer l'informatisation aux pratiques des juges, mais ce qu'elle a déjà pu changer. Il invite également à considérer une progression en plusieurs étapes de la numérisation de la Justice. L'intérêt de cette diversité d'acteurs pour la publication des décisions de justice reflètent la pluralité des enjeux qu'elle met à jour.

C'est dans ce contexte que se déroule le stage que j'ai mené au sein du projet de recherche e-Juris. Dans ce cadre, j'ai eu pour mission de travailler sur un corpus de décisions de justice des cours d'appels administratives et du Conseil d'État, pour en extraire des informations qui concernent les jurisprudences. Si nous avons vu ce terme précédemment, il nous faut ici le définir. Si les définitions de ce terme sont encore débattues, nous retiendrons celle développée par Évelyne Serverin : une jurisprudence est une interprétation de la loi sur un cas d'espèce, dont le statut est construit dans un processus de diffusion et de sélection des décisions de justice lié à la hiérarchie juridique ainsi qu'au travail de commentaires des juristes²². L'application de ces jurisprudences, est également liée au pouvoir normatif de la hiérarchie juridique, et en particulier du Conseil d'État. De plus la question du précédent juridique est plus importante au sein des juridictions administratives, qu'au sein de la justice judiciaire. La jurisprudence est donc un objet particulièrement intéressant à observer dans notre cas. Il s'agit donc de comprendre quelles informations sont utilisables au sein des décisions de justice, comment identifier et extraire les jurisprudences, ou encore de mettre en relation ces dernières avec des métadonnées afin de les caractériser. La question de recherche principale de ce projet est la suivante : Comment évoluent les jurisprudences dans le temps et dans l'espace ?

L'ensemble de ces réflexions nous invite à nous pencher sur les transformations qu'apportent avec elle la numérisation des documents de la justice administrative. Pour

²² Cette définition de la jurisprudence s'appuie sur les recherches d'Evelyne Serverin, et en particulier : Serverin, Evelyne. « Juridiction et jurisprudence : deux aspects des activités de justice ». *Droit et Société*, vol. 25, n° 1, 1993, p. 339-49.

cela, il est nécessaire de s'intéresser aux outils de Traitement Automatique de la Langue (TAL), de faire des analyses sémantiques et de chercher des spécificités et des ressemblances entre des documents textuels. Au travers de ce mémoire, nous montrerons les résultats de nos travaux au sein du projet de recherche E-Juris, pour donner à voir comment l'on peut observer et exploiter les décisions de justice sous leur format numérique, et comment le déploiement d'un ensemble de techniques et d'outils nous permettent d'objectiver des savoirs sur le fonctionnement de la justice administrative, et sur les jurisprudences de cette dernière. Il s'agit ici de savoir **comment la numérisation des documents juridiques permet de penser et de mettre en oeuvre des techniques de traitement de données afin de créer de nouvelles connaissances et de nouveaux savoirs sur le fonctionnement du Droit administratif et de ses institutions en France ?**

Nous verrons dans un premier temps quelles informations nous pouvons tirer des décisions de justice, et comment nous opérons une sélection des informations afin de constituer un corpus de documents exploitables. Ce travail suppose de comprendre la manière dont sont produites les décisions de justice, et de revenir rapidement sur l'histoire de la numérisation des documents administratifs. Cet examen, comme nous l'avons rapidement évoqué, est nécessaire pour redonner du sens aux données que l'on extrait des décisions. Dans un second temps, nous verrons comment il est possible de construire des analyses à partir de ces documents, d'en comprendre les limites, mais aussi les apports pour le Droit. Nous finirons par questionner les effets induits par l'exploitation des données, les évolutions que celle-ci peut amener, ou du moins les questionnements que se posent les acteurs sur ces évolutions. Nous répondrons finalement à notre problématique avec les éléments que nous avons développé tout au long de ce mémoire.

I. De la décision de justice aux données juridiques : constituer le corpus

La numérisation des documents juridiques est la première étape, sinon la plus importante, du développement d'outils informatiques, et d'études informatiques en Droit. Ces données sont nécessaires pour un certain nombre d'approches, et tout particulièrement pour les acteurs qui cherchent à développer des modèles d'intelligences artificielles. Ce travail repose cependant sur une forte annotation, sur un nombre d'objets souvent conséquents, et sur une quantité large de documents. Cette annotation n'étant pas réalisée sur les documents à disposition par l'administration, il faudrait ainsi annoter plus de 500.000 documents (pour ne prendre que les documents à notre disposition), qui comportent chacun un ensemble d'informations important. Ce lourd travail qui peut sembler au mieux fastidieux, au pire irréalisable, ne peut se réaliser qu'avec des moyens humains et matériels importants.

Notre approche ne repose donc pas sur la volonté de créer un modèle d'intelligence artificielle (IA) à même de faire des prédictions sur des affaires en cours, sur l'issue d'un procès, ou sur les sommes auxquelles peuvent prétendre les parties. Au contraire, nous avons choisi de travailler dans une perspective relativement inductive : nous avons souhaité voir quelles informations pouvaient être observées des données en présence, en mettant en lien des phrases spécifiques entre elles. Nous avons pu avancer sur cette piste grâce à la participation de juristes et de magistrats, car les documents ne prennent sens qu'à travers les usages qu'ils permettent, et les « rituels » qui permettent de les constituer. C'est aussi parce que le Droit et les documents de justices sont « encodés » juridiquement: *« Ce message, étroitement codé, est souvent riche d'informations sous-jacentes, qu'une interprétation directe ne permet pas à elle seule de révéler. Cela conduit parfois à procéder à une analyse plus subtile pour rechercher ce qui peut être latent derrière une formule, une expression, un mot ... Une analyse indirecte du message peut alors permettre d'atteindre par inférence et, au-delà de ce qui est écrit, ce que l'auteur a scellé ou a voulu sceller »*²³

Pour explorer ces sujets, nous souhaitons revenir dans un premier temps sur la production de la jurisprudence en France, et notamment par rapport à la production

²³ Frison-Roche, *Art. Cit*

juridique : le contentieux. Nous reviendrons sur le concept de jurisprudence et sur une histoire rapide de la recherche récente en Droit. Nous proposons ensuite de montrer les formes et le balisage que prennent les décisions de justice, et d'ainsi voir les informations que nous avons à notre disposition. Nous questionnerons également les approches que nous avons choisies d'utiliser (et celles que nous avons écartées) en termes de sélection, de mise en cohérence, des données. Mais avant de revenir sur ces questions liées à l'exploitation et la production des décisions de justice, nous revenons sur les possibilités permises par la numérisation de ces documents, la question de la jurisprudence, et les présupposés qui ont permis la création du stage.

A. Historique et questionnements autour de la numérisation des documents juridiques

Les liens entre l'informatique et le Droit, mais aussi entre la statistique et le Droit, ne sont pas aussi récents que le suppose la prolifération des discours actuels sur la justice prédictive, et particulièrement les risques qu'elle présente. La juridiction judiciaire française est dotée d'un appareillage statistique depuis le début du XIXe siècle : 1827 pour la justice pénale, 1831 pour la justice civile et commerciale²⁴. L'utilisation des informations statistiques n'est pas nouvelle dans l'administration française, et perdure depuis plus de 2 siècles, notamment pour appuyer les décisions, produire des effets de mesure. Cependant, ces statistiques ont longtemps été éloignées du travail scientifique en Droit, qui n'ont pas été considérées, et ne le sont pas toujours, comme des données pertinentes de recherche pour cette discipline. Cette exclusion relative tire ses origines des conceptions dont est issu la recherche en Droit, c'est-à-dire la manière dont les juristes considèrent le fonctionnement du Droit en France.

À l'opposé d'une conception *Common Law* du Droit, le Droit français ne se construit pas sur le principe du précédent obligatoire : on parle de Droit légiféré, ou romano-germanique. Cela implique avant tout que les juges, qui doivent appliquer la loi, ne sont pas obligés, *a priori*, d'appliquer les précédents dans les procès qu'ils mènent. Cependant : « *On sait, bien sûr, qu'il n'en est rien, et que les arrêts des cours suprêmes jouissent d'une probabilité de diffusion incomparablement supérieure à celle des*

²⁴ Serverin, Évelyne, *Art. Cit.*

décisions du fond »²⁵. Par cette citation, Évelyne Serverin nous invite à nous interroger sur la diffusion des décisions, des précédents, et par là, sur l'autorité que prennent les décisions de justice diffusées par les cours suprêmes : la Cour de Cassation et le Conseil d'État. En tant que cours suprêmes, elles jouent un rôle dans l'interprétation de la loi, en étant la dernière juridiction qu'il est possible de saisir, à l'exception des cours européennes dans certains cas. Cela signifie qu'elles rendent le dernier jugement possible, et qu'elles décident des orientations d'interprétation sur une loi, sur un litige : c'est ce qu'on nomme « attendu de principe » pour la justice judiciaire, et **considérant de principe** pour la justice administrative, et que l'on retrouve dans certaines décisions de justice des cours suprêmes. C'est également ce que l'on peut définir, dans le Droit français, comme une jurisprudence, c'est-à-dire la solution *généralement* donnée par les tribunaux à une question de Droit, prononcée par une juridiction supérieure dont les décisions sont connues des professionnels de justice, et que ces solutions sont suivies d'effet, au sens où les questions de Droit identiques seront globalement réglées de façon identiques. En ce sens, les juges bien qu'indépendant.es dans leur travail, puisqu'ils et elles ne sont pas contraint.es par le précédent, s'alignent très souvent sur les jurisprudences, quand bien même ils et elles contesteraient ces dernières. La hiérarchie de la justice impose une forme de contrôle du travail des juges, qui ne veulent pas forcément voir leurs décisions « casser » en appel. À défaut de hiérarchie sur un point de Droit, c'est également par routine, au vu de la grande production de la Justice, que certaines jurisprudences s'imposent.

En quoi ces considérations concernent-elles la recherche en Droit ? Parce que les juristes s'intéressent particulièrement à commenter les arrêts des cours supérieures pour définir les jurisprudences : « *plutôt que d'affirmer que seuls les arrêts de ce type peuvent légitimement faire jurisprudence, on peut dire qu'ils ont institutionnellement une vocation jurisprudentielle plus grande que toute autre décision juridictionnelle.* ». Ces arrêts, qui sont très largement minoritaires par rapport à l'ensemble des décisions de justice produit par l'administration juridique, bénéficient d'un regard tout particulier pour les juristes, qui les sélectionnent, les commentent, les diffusent, de même que ces arrêts ont bénéficié, et bénéficient encore, d'une large diffusion de la part des cours dont elles sont issues. Ces décisions sont les objets d'études de la plupart des juristes, qui relèguent les décisions qui ont moins de « vocations jurisprudentielles » au rang d'objets de peu de valeurs pour la recherche en Droit : « *Dans ce schéma vertical et hiérarchique, les millions de décisions*

²⁵ Serverin, Évelyne, *Art. Cit*

provenant des juridictions du fond n'ont pas de statut théorique, et ne présentent un intérêt que pour autant qu'elles participent à ce processus d'unification, en faisant apparaître des « précurseurs », des « résistances », ou des « divergences » de jurisprudence »²⁶.

Cette exclusion du champ de l'étude d'un vaste ensemble de décisions de justice, bien que ne présentant pas les pré-requis à une vocation jurisprudentielle, a plus récemment été au centre d'un regain d'intérêt par les acteurs de la recherche, et en particulier sous l'impulsion de Danièle Bourcier et la science du Droit. L'utilisation des techniques et des outils informatiques permet le développement de cet espace scientifique, notamment en mettant en avant les possibilités de travailler sur les décisions non plus pour étudier la « norme » - l'application de la règle - mais de prendre comme objet d'étude le Droit lui-même. Elle oppose de cette manière à la « Doctrine », que l'on désigne comme l'étude traditionnelle du Droit, à la science du Droit. Les approches qui suivront s'intéresseront plus particulièrement à l'analyse de la masse des décisions, notamment dans leur capacité à produire de la connaissance. Serge Bories et Marie-Anne Frisson-Roche, mais aussi Pascale Deumier, racontent comment l'on peut traiter le « contentieux », c'est-à-dire la masse des décisions « du fond ». Mais ces approches ne sont possibles que si les chercheur.ses ont un accès à ces décisions de justice dans des proportions conséquentes.

Comme le rappelle Marc Clément²⁷, et dans une autre mesure Évelyne Serverin, l'informatique est depuis longtemps un outil utilisé tous les jours par les magistrat.es et les professionnel.les de la justice, et notamment dans le processus de recherche et de sélection de l'information pertinente : en tant qu'outils d'aide à la décision. Serverin nous montre l'évolution de ce qu'elle appelle « l'espace jurisprudentiel »²⁸, c'est-à-dire l'ensemble des acteurs et des pratiques qui concourent à la construction et la diffusion des jurisprudences, au travers de l'adoption des outils informatiques par les institutions juridiques, ou la publication des documents de la justice : « *on peut constater que le premier niveau [L'utilisation de base de données et/ou d'outils pour appuyer l'expertise] correspond à l'évolution du métier du juge depuis quinze ans pour l'ensemble des types*

²⁶ Serverin, Évelyne. « De l'informatique juridique aux services de justice prédictive, la longue route de l'accès du public aux décisions de justice dématérialisées ». *Archives de philosophie du Droit*, vol. 60, n° 1, 2018, p. 23-47.

²⁷ Clément, Marc, « Algorithmes au service du juge administratif : peut-on en rester maître ? », *AJDA*, 2017 p.2453

²⁸ Jeammaud, Antoine, Serverin, Evelyne, « *Concevoir l'espace jurisprudentiel* », *RTD civ.*, 1993, p. 91

de contentieux. En effet, l'introduction de bases de données jurisprudentielles depuis les années 2000 a considérablement modifié le travail du juge. Par ailleurs, aux Etats-Unis comme dans d'autres pays, la justice pénale utilise des outils statistiques pour apporter une expertise au juge dans ce qui est appelé l'Evidence-Based Sentencing. Enfin, les échanges dématérialisés entre les juridictions et les auxiliaires de justice sont une réalité qui s'est rapidement installée dans la juridiction administrative »²⁹. Cette évolution³⁰ trouve son aboutissement aujourd'hui avec la loi du 7 octobre 2016 pour une République numérique - ou loi Lemaire - qui introduit la publication par défaut des données publiques. En affirmant ce principe, les chercheurs se retrouvent alors avec une quantité importante de décisions de justice numérisées à leur disposition, à même d'être traitées statistiquement.

Il nous faut alors comprendre comment ces documents sont produits, lesquels sont stockés sur des bases de données, et quelles informations peuvent en être extraits : « Ainsi le recueil et la constitution des corpus doivent-ils prendre en compte les spécificités textuelles, génériques et discursives des matériaux considérés, car ces spécificités en déterminent les régimes d'interprétation (par exemple, à propos du tweet politique, voir Longhi, 2013) »³¹.

B. Les décisions de justice : de la décision aux données.

1. La base de données JADE et les décisions de justice administrative

L'ensemble des documents avec lesquels nous travaillons sont disponibles publiquement sur le site de la Direction de l'Information Légale et Administrative (DILA), au sein de la base JADE. Cette dernière met à disposition plusieurs « types » de décisions de justice administrative³² : Celles du Conseil d'État, celles des cours d'appels

²⁹ Clément, Marc, *Art. Cit.*

³⁰ Que l'on retrouve décrite, par exemple, dans les articles de Danièle Bourcier et Primavera De Filippi, ou celui d'Évelyne Serverin.

³¹ Longhi, Julien. « Humanités, numérique : des corpus au sens, du sens aux corpus ». *Questions de communication*, vol. 31, n° 1, septembre 2017, p. 7-17.

³² Que l'on retrouve décrites sur le site de la DILA : https://echanges.dila.gouv.fr/OPENDATA/JADE/%20DILA_JADE_Presentation_20171212.pdf

administratives, et certaines provenant des tribunaux administratifs. Si cet échantillon essaye d'être le plus exhaustif possible par rapport à l'ensemble des décisions depuis la création de ces institutions juridiques, nous n'en avons qu'une partie limitée. Pour le Conseil d'Etat, nous avons : les « grands arrêts » de la justice administrative³³ depuis sa création, les décisions publiées au recueil Lebon depuis 1965, ainsi qu'une sélection limitée de décisions qui ne sont pas publiées au recueil depuis 1975. Pour les cours d'appels, nous avons une sélection pour chacune des 8 cours d'appels, depuis leur création, c'est-à-dire depuis 1989 pour les plus anciennes. Pour les tribunaux, nous avons une petite sélection depuis 1965, qui sont toutes des décisions qui ont été retenues au recueil Lebon.

Au total, cette base de données comporte 491506 fichiers au moment où nous avons constitué notre corpus : 161621 provenant du Conseil d'État, de la Cour de discipline budgétaire et financière et du tribunal des conflits, 322101 des cours d'appels administratives et 6515 des tribunaux administratifs. 1710 décisions proviennent du tribunal des conflits et 68 de la Cour de discipline budgétaire et financière. La plus ancienne décision date de 1873, les plus récentes de février 2022 et la base de données est régulièrement mise à jour. Une décision mal annotée comporte la date 2999. La grande majorité des décisions de justice se situent entre 1965 et 2022³⁴. Nous reviendrons plus tard sur les processus de sélection des données³⁵.

Pour comprendre de quelle manière ces documents sont exploitables, il nous faut d'abord revenir sur le processus de construction d'une décision de justice. Tout d'abord, revenons sur la distinction entre la justice administrative et la justice judiciaire : « *en France, nous l'avons dit, il existe depuis Napoléon deux systèmes juridiques entièrement différents : le judiciaire et l'administratif. Le premier règle les disputes des particuliers entre eux — le Droit privé — ainsi que des crimes ou des délits — le Droit pénal ; le second connaît de tous les conflits avec l'administration* »³⁶. Cette distinction implique que pour saisir la justice administrative, le « requérant » doit s'opposer à un fait ou à un acte de l'administration. Par exemple, une personne ou une entité peut se retourner contre l'administration aussi bien pour un titre de séjour refusé (ou approuvé), que pour

³³ On peut parfois les retrouver sous la dénomination « GAJA ».

³⁴ Voir Annexe 1.1 : Répartitions par années des décisions de justice administrative

³⁵ Voir 1.C : Que faire de ces données ?

³⁶ Latour, Bruno. « 1. Sous l'ombre de Bonaparte ». *La fabrique du Droit*, La Découverte, 2004, p. 11-81. *Cairn.info*, <https://www.cairn.info/la-fabrique-du-Droit--9782707144720-p-11.htm>.

les travaux sur une place publique. Le requérant n'a pas à être directement concerné par les actions de l'administration, il doit cependant pouvoir *motiver* sa demande. C'est de cette manière que l'on accède au tribunal administratif, que l'on peut aussi nommer tribunal de 1^{ère} instance, ou encore « juridiction du fond ». Une fois la décision rendue³⁷, les parties peuvent faire appel de la décision, ou se « pourvoir en cassation » s'ils ont déjà fait appel. C'est de cette manière que, respectivement, l'affaire se poursuit en cour d'appel ou au Conseil d'État.

Ce mouvement hiérarchique - en plus de créer un phénomène de contrainte relative comme nous avons pu le voir précédemment - est particulièrement important pour comprendre certains aspects des analyses que nous pouvons produire, étant donné qu'elles proviennent presque exclusivement des juridictions d'appels et supérieures. Nous avons accès à des décisions de tribunaux mais dans des proportions très faibles. Cela implique que nous n'avons pas accès à une grande partie du *contentieux* traités par les tribunaux, soit les décisions des affaires qui n'accèdent pas aux cours d'appels ou au Conseil d'État. Comme le rappelle Pascale Deumier : « *Il existe des millions de décisions de justice rendues par les juges du fond qui, faute d'être connues hors des parties au litige, ne peuvent produire d'effet normatif au-delà de cette sphère limitée. La justice prédictive, comme outil d'exploitation de cette masse de décision, permet de les porter à la connaissance et, dans le même mouvement, d'indiquer un effet normatif.* »³⁸. Nous ne pouvons étudier les potentiels normatifs de ces décisions, comme c'est également le cas pour les juridictions supérieures, qui *dans une certaine mesure* n'ont pas forcément de visibilité, et donc de contraintes, sur les jugements rendus par les juridictions du fond.

Pour revenir sur les décisions de justice à notre disposition, il est important de comprendre comment elles sont produites afin de saisir les informations que nous pouvons extraire de celles-ci. Il s'agit donc de comprendre quels objets nous pouvons construire à partir des informations balisées qui sont présentes dans les fichiers XML des décisions.

³⁷ Pour comprendre comment la décision est prise et rédigée, Bruno Latour en donne une description précise dans le premier chapitre « *1. Sous l'ombre de Bonaparte* », de son ouvrage *La fabrique du Droit*.

³⁸ Deumier, Pascale, *Art. Cit.*

2. Les métadonnées

Tout d'abord, l'ensemble des fichiers récupérés sont au format XML, les données sont donc balisées, et accessibles automatiquement avec des outils de programmation. Nous voulons voir ici quelles données sont accessibles systématiquement et de manière à pouvoir automatiser la collecte des données. Pour cela, nous nous appuyons sur les balises présentes au sein du fichier, dont on peut retrouver un exemple en annexe³⁹.

Il y a tout d'abord 3 principales balises, qui créent trois « blocs » d'informations sensiblement différentes que nous allons décrire. Gardons également à l'esprit que les documents provenant du Conseil d'État sont légèrement différents de ceux des cours d'appels : nous reviendrons rapidement sur cette différence. En premier lieu, nous avons la balise <META>, qui regroupe l'ensemble des métadonnées associées à la décision. Elles sont divisées en 2 sous-balises : <META_COMMUN> et <META_SPEC>. La première regroupe des informations générales sur le dossier, notamment son nom de fichier présent, et son identifiant précédent, c'est-à-dire avant que le dossier arrive en appel ou au Conseil d'État. La balise <ORIGINE> nous permet de savoir quelle juridiction à émis cette décision de justice au sein de la base JADE. Les autres balises concernent la nature du document, ici du texte, et une URL, qui semble être la position du fichier au sein de la base JADE. Au sein de la balise <META_SPEC>, on retrouve les informations qui portent directement sur la décision : le titre, sa date de publication, la juridiction qui l'a rédigée, le numéro du dossier, la solution retenue par cette juridiction. On parle de solution pour évoquer l'issue du litige. Elle peut prendre plusieurs formes, mais très souvent, c'est le terme REJET qui revient, soit à la faveur du requérant, soit pour l'administration visée par le litige. Nous verrons plus tard pourquoi nous n'utilisons pas cette variable dans nos futures analyses.

La balise <META_SPEC> comprend également d'autres métadonnées qui peuvent nous intéresser. D'abord la formation de jugement, c'est-à-dire, la manière dont les magistrats se sont réunis pour prononcer le jugement. Cette information nous renseigne en particulier sur l'importance de la décision : en effet, plus la formation serait large, c'est-à-dire avec d'autres « chambres », plus cela signifie que la décision pose des questions en matière d'interprétation de la loi, et donc en matière jurisprudentielle. Sans revenir sur l'organisation du Conseil d'État, qui n'est pas le propos de notre sujet, et qui,

³⁹ Voir annexe 1.2. Exemple d'une décision de justice

par ailleurs, a été très bien décrite par Bruno Latour⁴⁰, nous pouvons dire que dans la décision décrite en annexe, la formation s'est faite entre les sous-sections 8 et 9 du Conseil d'État, ce qui n'est pas une décision sans importance, mais dont l'intérêt est relatif. On retrouve, juste après, une balise qui qualifie le type de recours fait devant le tribunal, c'est-à-dire le sujet, en termes de procédure, du litige : par exemple, est-ce que celui porte plutôt sur un « excès de pouvoir » de la part de l'administration ? Ou sommes-nous plutôt dans du « plein contentieux fiscal » ? Quand on parle de recours pour excès de pouvoir, on évoque les litiges qui concerne la régularité juridique d'un acte administratif (ou de la personne), c'est-à-dire de savoir s'ils sont légaux ou non. Dans le cas du plein contentieux, le ou la requérant.e faire valoir ses Droits auprès du juge, que l'administration n'aurait pas appliqué, ou pas pris en compte. Ces types de recours déterminent également les possibilités du juge en terme de décision : dans le cas de l'excès de pouvoir, le juge ne peut que prononcer l'illégalité ou non d'un acte administratif, ce n'est pas le cas pour les recours en plein contentieux. Ces deux catégories sont les plus courantes, mais il existe un grand nombre de modalités à cette variable : pour n'en citer que quelques-unes : les référés, les rectifications pour erreur matérielle, les tierces oppositions,...

Pour terminer avec la balise <META_SPEC>, il reste les noms des parties en présence au litige, qui sont enlevés des fichiers par souci d'anonymisation, ainsi que les noms du président, du rapporteur public, du commissaire du gouvernement, et ceux des avocats dans certains cas. Enfin, nous avons également la mention du code de publication au recueil Lebon. Cette variable nous intéresse particulièrement pour plusieurs raisons : le recueil Lebon est un ouvrage, régulièrement mis à jour, qui publie les arrêts⁴¹ du Conseil d'État, des cours d'appels, et des tribunaux administratifs. Cependant ce ne sont pas n'importe quels arrêts qui sont publiés au sein du recueil, mais ceux qui sont les plus « importants », c'est-à-dire qui sont les plus à même d'influencer la règle de Droit. On dit qu'ils sont « publiés au recueil Lebon ». Mais un grand nombre de ces décisions est publié au recueil Lebon, et toutes n'ont pas la même importance, comme celles qui sont « inédites au recueil Lebon ». C'est ainsi qu'il existe un code, qui permet d'évaluer l'importance des arrêts. Les plus importantes, et celles qui sont le plus à même, d'après les juristes, d'influencer la jurisprudence, sont les arrêts publiés en « A », suivis de ceux publiés en « B ». La qualification dans ces deux catégories est la prérogative du

⁴⁰ Latour, Bruno, *Op Cit.*

⁴¹ Une décision de justice peut également être appelé un arrêt.

Conseil d'État, ce qui signifie que les cours d'appels ne peuvent attribuer cette importance à une décision. Elles peuvent par contre indiquer cette importance en qualifiant une décision en « C+ », à l'inverse des décisions publiées en « C », qui sont d'une importance minimale. Cette logique de diffusion, de qualification, et finalement de contrôle de la jurisprudence est proche de ce que décrit Évelyne Serverin sur les mécanismes à l'origine de la promotion de la jurisprudence⁴².

3. Le texte de la décision

Nous avons vu les métadonnées au sein de la balise <META>, nous nous intéressons maintenant au cœur de notre sujet, et à la balise <TEXTE>. Nous y retrouvons, pour les décisions du Conseil d'État, deux sous-balises : la balise <BLOC_TEXTUEL>, qui comprend la balise <CONTENU>, et la balise <SOMMAIRE>, qui contient des balises <SCT> et <ANA>. Les arrêts des cours d'appels ont généralement une balise <SOMMAIRE> très succincte, et pas de balise <ANA>. La balise <CONTENU> représente l'ensemble du texte de la décision de justice, que l'on peut retrouver dans la décision au format papier⁴³. Ce texte a pour objectif de communiquer la décision de justice vers un ensemble d'acteurs et d'actrices : d'abord les parties du litige, qui peuvent y lire la décision du juge et de la formation de jugement. Également les observateurs « extérieurs » au litige, comme des journalistes, ou des citoyens qui souhaitent s'informer sur ce procès. Enfin, et comme le fait remarquer Serge Bories et Anne-Marie Frison-Roche, les juges entre eux sont observateurs des décisions de justice : « *les parties au procès ne sont finalement les récepteurs que de la solution donnée au litige, alors que les destinataires du discours judiciaire qui s'incarne dans la décision sont les autres juges. Ainsi, la décision de justice a une double nature : elle est à la fois un message de type « instrumental », c'est-à-dire destiné à produire un certain effet sur le récepteur, en même temps qu'un message de type « représentatif », c'est-à-dire mettant en scène l'émetteur et révélant notamment son état d'esprit* »⁴⁴. Il existe ainsi plusieurs « niveaux » de lecture au sein d'une décision, qui reposent sur une syntaxe et une construction particulière du texte : un encodage spécifique⁴⁵.

⁴² Serverin, Évelyne, « Juridiction et jurisprudence : deux aspects des activités de justice », *Art. Cit.*

⁴³ Voir annexe 1.3 : Exemple d'une décision de justice au format papier

⁴⁴ Frison-Roche, Marie-Anne, *Art. Cit.*

⁴⁵ *Ibid.*

L'écriture d'une décision repose sur des conventions que l'on peut retrouver dans la forme et par des documents produits par le Conseil d'État à l'attention des magistrats, pour rédiger celle-ci. On retrouve 3 parties principales dans ces décisions : les **visas**, qui contiennent les éléments de la procédure du procès en cours, les arguments de chacun des parties, et les références légales mobilisées. Ils visent à « *présenter le cadre du litige et permettent, en particulier, aux parties de s'assurer que le juge a correctement et complètement analysé et restitué leurs conclusions et leur argumentation* »⁴⁶. On y expose ainsi les **conclusions** avancées par les parties, c'est-à-dire les demandes des parties, argumenté par des **moyens**, c'est-à-dire la réflexion juridique présentée par les parties au sein de **mémoires**.

On retrouve ensuite les **motifs**, qui sont les explications du sens de la décision par le juge : « *Ils permettent aux parties de comprendre les raisons qui ont conduit le juge à rejeter ou à faire Droit à leurs prétentions. Ils éclairent également la communauté des juristes et des membres de la juridiction administrative sur la place que prend la décision dans la jurisprudence qu'elle reprend ou qu'elle modifie. L'équilibre délicat recherché par le juge dans la rédaction de ses décisions se cristallise ainsi dans ses motifs.* »⁴⁷. On retrouve dans cette citation l'idée d'un double discours, lié à la multiplicité des acteurs à même de lire, pour des raisons différentes, la décision de justice. Le paragraphe qui introduit les motifs commencent dans la très grande majorité des cas par la phrase « *Considérant ce qui suit* », et est un enchaînement de **considérants**, qui examine et statue sur chacune des conclusions. Le juge est obligé de donner une réponse pour chaque conclusion présentée par les parties. On se retrouve donc avec un enchaînement de phrase commençant par « *Considérant que [...]* », qui correspondent aux considérants, et de phrase commençant par « *Sur [...]* », qui marque l'ouverture d'une réponse sur une conclusion. C'est dans les motifs que le juge reprend la jurisprudence et donne du sens à la décision.

Les motifs sont suivis, enfin, par le **dispositif**, qui présente les décisions qui découlent des motifs, en un ou plusieurs articles. Les articles doivent reprendre au mieux la structure du jugement formulée précédemment par le juge, afin d'isoler chaque effets liés aux réponses aux conclusions. De cette manière, les décisions sont plus facilement

⁴⁶ Cette citation est tiré directement d'un document du Conseil d'État à l'attention des juges pour les aider et les guider dans la rédaction des décisions de justice administrative. L'autorisation de le diffuser n'a pas été donnée dans la cadre de ce mémoire.

⁴⁷ *Ibid.*

identifiables pour les lecteurs de la décision. La balise <CONTENU>, à ces égards, regroupe tous les éléments de langage et de réflexions qui constituent la décision de justice.

Enfin, si la balise <SOMMAIRE> est relativement peu importante pour les décisions des cours d'appels, elle l'est beaucoup plus pour les arrêts du Conseil d'État, dans la mesure où les balises <ANA> présentent des éléments particulièrement importants. En effet, ces balises reprennent certains des considérants, ou des re-formulations des considérants présents au sein des motifs, pour isoler un point important de la décision. Ces paragraphes particuliers, que l'on appelle **considérant de principe**⁴⁸, sont les points de Droit qui sont mis en avant par la formation de jugement du Conseil d'État : en somme il s'agit des points de Droit qui doivent faire jurisprudences au sein de la justice administrative. En plus des considérants de principe, on retrouve au début de chacun de ces paragraphes, le **code PCJA** auquel se réfère ceux-ci, c'est-à-dire un code qui permet d'identifier le type de contentieux auquel s'applique la jurisprudence nouvellement énoncée. La nomenclature du Plan de Classement de la Jurisprudence Administrative se décline en 62 entrées, qui elles-mêmes se déclinent en un nombre variable d'entrées. Un code peut comporter jusqu'à 6 « niveaux », c'est-à-dire 6 enchaînements de nombre, qui expriment un contentieux spécifique sur lequel s'applique la jurisprudence. Au premier niveau on peut retrouver des entrées telles que : « Procédure », « Étrangers » ou « Contribution et taxes », qui donne une idée plus ou moins précise des litiges.

C'est ainsi avec l'ensemble de ces éléments que nous allons par la suite travailler pour analyser l'évolution des jurisprudences. Il faut ici détailler les données exploitables, et dans quelles mesures nous pouvons les utiliser dans des perspectives statistiques.

⁴⁸ En voici un exemple : « 68-025-02 La circonstance que les documents produits à l'appui d'un dossier de demande de certificat d'urbanisme seraient insuffisants, imprécis ou comporteraient des inexactitudes, n'est susceptible d'entacher d'illégalité le certificat d'urbanisme qui a été accordé que dans le cas où ces omissions, inexactitudes ou insuffisances ont été de nature à fausser l'appréciation portée par l'autorité administrative sur la conformité du projet à la réglementation applicable. »

C. Comment utiliser ces données ? Sélectionner l'information

1. Sur les métadonnées

Si l'on voit assez rapidement dans quelle mesure on peut utiliser les métadonnées associées à un fichier, il faut voir si elles sont exploitables, c'est-à-dire suffisamment bien encodées, ou même présentes, dans les fichiers. En parallèle, il est plus difficile de voir dans quelle mesure le texte est exploitable, et comment en extraire des données. La codification, généralement rigoureuse, des décisions de justice peut en ce sens nous aider à effectuer des traitements automatiques du langage sur les textes : « *Le Droit comme activité intellectuelle bénéficiant d'un formalisme marqué est, en effet, un terrain favorable à l'application des techniques associées à l'intelligence artificielle. Les structures des textes juridiques sont des points d'appui précieux pour construire une analyse automatique. On peut ainsi sans trop de difficulté identifier des références à des articles de codes ou de lois ou à des jurisprudences, ce qui place d'emblée un texte juridique dans un réseau d'autres textes* »⁴⁹.

En ce qui concerne les principales métadonnées, les dates peuvent être très utiles pour étudier la question de l'évolution temporelle des jurisprudences, notamment en observant les considérants de principe au cours des années. L'année semble une information utile en ce sens, ce qui est peut être moins le cas du mois et le jour de la décision. Il n'est pas particulièrement utile à l'analyse d'avoir ces détails dans la mesure où cela donne une trop grande finesse pour observer des phénomènes sur un éventail de 150 années. De plus, il n'est pas particulièrement utile de savoir, au sein d'une même année, les temporalités des décisions de justice, on peut s'arrêter à savoir quelles décisions ont été prises et rédigées la même année. Comme nous avons pu le voir dans l'annexe de la répartition des décisions par année⁵⁰, on observe un fort déséquilibre entre les documents avant 1965, ceux entre 1965 et la fin des années 1980, et ceux de 1990 à 2022. On peut essayer de savoir combien de documents nous avons pour chacune de ces périodes : avant 1965 il y a 1261 décisions, entre 1965 et 1986 il y a 25672 décisions, et il y a 464572 décisions écrites (et numérisées) à partir de 1986. Cette évolution peut s'expliquer par plusieurs choses, d'abord la création en 1989 des cours d'appels, qui ont

⁴⁹ Clément, Marc, *Art. Cit.*

⁵⁰ Voir annexe 1.1 : Répartition des décisions de justice par année

probablement augmenté la production de décisions, mais également la numérisation des documents. Il existe très peu de documents numérisés avant 1965, et c'est pour ces raisons que nous n'en tiendrons pas compte dans l'analyse pour des soucis de représentations des données. De plus, il est peu probable que des décisions rédigées à la fin du XIXe siècle soit reliées à des décisions récentes, d'abord parce que la loi a beaucoup évolué (par exemple avec la prise en compte du Droit communautaire), ainsi que la jurisprudence.

Il nous paraît également important de conserver la variable qui précise la juridiction qui a pris la décision, notamment pour voir si certaines juridictions adoptent plus souvent certaines décisions que d'autres, où si elles sont plus concernées par un contentieux en particulier. Comme nous avons pu le voir précédemment, nous avons quelques décisions des tribunaux (6515), et principalement des cours d'appels et du Conseil d'État. Nous préférons garder ici tous les documents, dans la mesure où les décisions des tribunaux ont été sélectionnées par le Conseil d'État, et ont été retenues pour publication ou mention au recueil Lebon. Elles ont une valeur jurisprudentielle aux yeux des conseillers d'État. En plus des tribunaux, des cours d'appels et du Conseil d'État, nous retrouvons également des décisions qui émanent du tribunal des conflits et de la cour de discipline budgétaire et financière. Le premier permet de régler les questions de compétences entre la justice administrative et la justice judiciaire sur un dossier : quand il y a un doute qui réside sur la juridiction qui doit être saisie par le requérant, le tribunal des conflits est à même de juger, et d'attribuer un dossier à une juridiction⁵¹. La seconde sanctionne « *les atteintes aux règles régissant les finances publiques commises par toute personne intervenant dans la gestion des affaires publiques* »⁵². Elle est compétente pour juger les fautes de gestion de l'argent public, et notamment des responsables d'entreprises publiques. Elle n'est cependant pas compétente pour juger les actes personnellement signés par les ministres, qui sont contrôlés par l'Assemblée Nationale. Même si cette dernière ne représente que 68 décisions, et que le tribunal des conflits n'en représente que 1710, nous les conservons pour voir si ces juridictions reviennent dans certaines

⁵¹ « *il surgit forcément de temps à autre des affaires dont on ne sait pas si elles appartiennent au judiciaire ou à l'administratif. Le malheureux justiciable risque par conséquent de ne pas savoir à quel tribunal se vouer si chacun de ceux qu'il a saisis se déclare à son tour incompetent. C'est pour éviter ce double bind que l'on a créé en 1849 un tribunal spécial dont le rôle n'est pas de trancher les affaires au fond, mais de décider laquelle des deux branches du Droit sera compétente pour les entendre.* ». Latour Bruno, *Op. Cit.*

⁵² « Cour de discipline budgétaire et financière ». *Wikipédia*, 2 octobre 2021. *Wikipédia*, https://fr.wikipedia.org/w/index.php?title=Cour_de_discipline_budg%C3%A9taire_et_financi%C3%A8re&oldid=186811489.

jurisprudences. Ainsi, nous conservons ici toutes les informations disponibles. Pour conserver ces données, il nous a fallu re-coder quelques noms de juridiction qui avait été écrites de différentes manières. Nous avons donc remplacé toutes les mentions des mots « Cour d'appel administrative » par « caa », et nous avons converti toutes les lettres en minuscule pour éviter les différences de « cases » entre les mots.

En ce qui concerne les types de recours, on peut s'intéresser à la présence des jurisprudences des recours, ou de voir, de manière générale, quel type de recours revient le plus souvent au sein de la justice administrative. De la même manière que pour les juridictions, une phase de re-codage est nécessaire : d'abord parce qu'il existe 242 types de recours différents dans les métadonnées, et que pour 195 des types de recours, il y a au mieux 10 arrêts associées, ce qui fait qu'au maximum, 1950 décisions ne font pas parties des 57 types de recours les plus mentionnés, c'est-à-dire 0,3% des décisions. De plus, une partie des types de recours peut être simplifié : « contentieux fiscal » et « fiscal » sont les mêmes catégories que « plein contentieux fiscal », ou encore « recours pour excès de pouvoir plein contentieux » est associé à un recours pour « excès de pouvoir ». Une partie conséquente des décisions n'a pas non plus de type de recours mentionné (19,8%), mais plus de 80% des décisions ont une mention, ce qui permet de pouvoir utiliser cette variable. Pour au mieux pouvoir l'utiliser, nous avons donc re-coder les types de recours, et intégrer tous les types de recours trop peu utilisés aux décisions sans mention d'un recours⁵³. On constate également que la majeure partie des décisions qui ne présentent pas de mention d'un type de recours sont délimitées temporellement entre 1986 et 2002⁵⁴. Cela ne nous empêche pas d'utiliser cette variable, d'autant plus que le nombre de décision est plus faible entre 1986 et 2002 qu'après 2002.

Il n'en est pas de même pour les informations contenues dans la balise <SOLUTION>. En effet, s'il pourrait être particulièrement intéressant de voir si des jurisprudences ont des tendances pour certaines solutions, on remarque que 82,2% des décisions n'ont pas de mention d'une solution. Seul 9,7% ont la mention « rejet », qui est la mention que l'on retrouve le plus souvent dans nos fichiers, suivi par la mention « satisfaction totale » à 1,1% seulement. De plus, il existe 2044 modalités différentes, dont 1858 mentions sont associées à moins de 10 décisions de justice : cela représente, au mieux, 18580, c'est-à-dire, 3,7% des décisions. Le fait qu'il y ait trop peu de décisions

⁵³ Voir annexe 1.4 : Graphique des types de recours qui sont mentionnés plus de 50 fois au sein des décisions de justice

⁵⁴ Voir annexe 1.5 Graphique des types de recours par année sur les recours mentionnés plus de 50 fois au sein des décisions de justice

avec une mention d'une solution ne nous permet pas d'utiliser cette potentielle variable. Nous n'utilisons pas non plus les informations qui concernent les parties (si elles sont présentes), les avocats et les magistrats, d'abord pour des raisons d'anonymisation, d'éthiques, et de droit : depuis le 23 mars 2019⁵⁵, le profilage des magistrats est interdit par la loi, ce qui implique que tout traitement statistique associés à l'identité d'un magistrat est interdit.

Également nous ne prenons pas en compte la métadonnée qui concerne la formation de jugement, dans la mesure où celle-ci est différente en fonction des juridictions, et des tribunaux. De plus, on peut avoir une information sur l'importance de la décision de justice grâce à l'indice de publication au recueil Lebon. La très grande majorité des décisions sont publiés en « C » (87,9%). 7,3% des décisions sont publiées en « B » et 4,5% des décisions sont publiées en « A ». On retrouve également des décisions publiées en « R », et en « C+ », mais qui sont très peu significatives (respectivement 0,09% et 0,08%). Enfin, on conserve le nom du fichier, qui nous permet d'avoir un identifiant fiable avec ces métadonnées pour pouvoir faire des jointures sur les données liées aux textes des arrêts.

2. Sur le texte des décisions

Comme nous l'avons vu précédemment, on retrouve trois parties distinctes dans le texte : les visas, les motifs, et les dispositifs. Chacune de ces parties sont délimitées par des phrases et des mots spécifiques. Les visas commencent par le mot « Vu », et se terminent au début des motifs, avec le mot « considérant », qui se terminent eux-mêmes par le début du dispositif, avec le mot « DÉCIDE : ». Dans les trois cas de figure, il a fallu en parti re-coder certaines déclinaisons différentes de ces « mots-frontières » : Pour rendre l'analyse de texte automatisable, il a fallu identifier les spécificités de chacune de ces parties, et réduire au mieux le nombre de cas particulier.

Commençons par les visas. Comme nous l'avons montré précédemment, ces derniers servent à montrer l'état du procès, les différentes conclusions présentées par les parties, et les références légales qui sont mobilisées dans le cadre du litige. Chacun de ces aspects est présentée dans une partie spécifique des visas, que l'on peut récupérer automatiquement grâce à de nouveaux « mots-frontières » : en cours d'appel, on retrouve très souvent 4 parties délimitées par les phrases : « Vu [...]», « Procédure devant la

⁵⁵ Loi du 23 mars 2019, de programmation 2018-2022 et de réforme pour la justice.

Cour », « [le.la.les requérant.es] soutien[t/nent] que » , « Vu les autres pièces du dossier ». Au Conseil d'Etat, on ne retrouve que deux parties bien distinctes : « Vu », « Vu les autres pièces du dossier ». Si l'on prend en compte les déclinaisons plus ou moins larges de ces phrases, on trouve des visas dans 91.4% des décisions de justice, c'est-à-dire 449378 fichiers sur 491504. Ce découpage en parties nous permet de récupérer des informations spécifiques au sein des visas, et en particulier ici les références légales, qui peuvent nous donner des informations sur le contentieux, ou qui peuvent nous permettre par la suite d'observer la diversité d'application d'une jurisprudence.

En revanche, nous n'utilisons pas les informations que l'on peut tirer des dispositifs. Ces derniers ne nous donnent pas beaucoup plus d'informations que les décisions prises par le magistrat à l'issue du procès que l'on retrouve dans les métadonnées. Dans une certaine mesure, cela pourrait nous permettre de trouver la solution apportée par le ou la juge quand nous ne l'avons pas présentée dans les métadonnées. Mais la difficulté à identifier automatiquement la partie pour qui la décision a été prise, ou de savoir sur quel objet porte la décision, rend difficile l'identification des solutions. Dans l'exemple que l'on peut retrouver en annexe 1.2, le juge décide que 1. La requête de Mr. X est rejetée et 2. que la décision sera transmise au ministre. Ces décisions nous permettent de voir qu'il y a rejet de la requête, mais sur quoi porte la requête ? Mr. X est-il le requérant en première instance ou en appel ? En croisant ces informations avec les visas, on pourrait avoir un peu plus de précision, mais cela nécessite un constant travail de vérification, et possiblement d'annotation, pour comprendre la portée de la décision.

La partie qui nous intéresse le plus est celle des motifs, et ce pour plusieurs raisons : d'abord parce que c'est l'endroit dans la décision où le ou la juge dévoile le déroulé de sa pensée qui lui amène à prendre des décisions. Si on retrouve les actes pris au sein du dispositif, on retrouve également les décisions prises au sein des motifs, et argumentées avec une réflexion juridique, qui reprend aussi les règles de Droit et la jurisprudence de la justice administrative. Et c'est particulièrement ce dernier point qui nous intéresse. Dans la mesure où, au sein du Droit français, la référence à la jurisprudence n'est pas explicite, il est très difficile pour un oeil non expert du Droit, et en particulier du Droit administratif, de repérer les références à la jurisprudence au sein de la décision. C'est précisément parce que ces références sont des phrases particulières, qui sont très souvent des reprises quasi exactes des considérants de principe mis en avant par le Conseil d'État, que l'on s'intéresse aux motifs. C'est donc en traitant

automatiquement les considérants des décisions de justice, en les mettant en relation avec les considérants de principe des décisions du Conseil d'État que l'on peut essayer de voir dans quelle mesure les jurisprudences évoluent dans le temps, mais aussi dans l'espace juridictionnel, entre autre.

Un ensemble de pré-traitement est d'abord nécessaire : il faut d'abord découper chaque considérant en document distinct, afin d'avoir non pas des phrases, mais bien des paragraphes qui correspondent à chaque considérant de notre corpus de décision. Ce travail se fait en particulier sur les balises
 qui délimitent chaque paragraphe, mais également avec les triples espaces ou plus, qui marquent la fin d'un paragraphe. Sur l'ensemble des décisions de justice, on récupère des considérants dans 93.4% des cas, c'est-à-dire sur 459291 décisions. Sur celles-ci, on retrouve 5.0107.48 considérants desquels nous n'avons pas soustrait les titres des conclusions. Ce traitement réalisé, il nous faut également récupérer les considérants de principe, que l'on retrouve dans les décisions du Conseil d'État, mais pas n'importe lesquelles. Nous ne récupérons pas les décisions publiées en « C » au recueil Lebon, dans la mesure où ces dernières ne sont pas considérées comme importantes par l'institution, ce qui amène également à penser qu'elles ne sont que peu susceptibles d'être ré-employées par les tribunaux et les cours d'appel. Sur l'ensemble des fichiers provenant du Conseil d'État (161092), on récupère 73151 considérants de principe fichés en « A » ou en « B » au recueil Lebon. En plus du considérant de principe, nous récupérons également le ou les codes PCJA qui lui sont associés, et qui nous permettent d'identifier le ou les types de contentieux sur lesquels portent le considérant de principe. On récupère ces codes au début des paragraphes, et on peut constituer un tableau qui indexe les considérants de principe avec leurs codes PCJA.

Ainsi, on récupère un certain nombre d'informations du texte lui-même, et c'est dans la mise en relation des documents nouvellement constitués que nous pouvons également tirer de l'information. C'est également en croisant ces relations avec les variables que nous avons mises à jour que nous pouvons analyser des comportements sur les individus qui nous intéressent ici : les jurisprudences.

II. Des données juridiques à la science du Droit : l'apport du numérique et le poids des choix

La constitution du corpus, bien que n'étant pas entièrement inductive, ne définit pas les résultats à l'avance. Et c'est bien dans cette perspective que nous souhaitons mener nos analyses. Pour interroger l'évolution des jurisprudences à l'aune des variables que nous avons pu constituer, il s'agit d'abord de mettre en lien ces considérants de principe du Conseil d'État avec les considérants des décisions de justice qui les reprennent. Ce travail est permis notamment par la conversion des phrases en vecteurs interprétables mathématiquement, et qui constituent des coordonnées dans un espace de mots. Ce travail suppose d'abord un second traitement sur les considérants (des décisions ou de principe) afin de rendre le plus comparable possible les documents, sans pour autant en biaiser substantiellement la comparaison. Il s'agit donc dans un premier temps de voir comment traiter les phrases afin de justifier nos choix des techniques de conversion vectorielle plutôt que d'autres. C'est ensuite en comparant ces vecteurs entre eux à l'aide du cosinus que l'on peut observer s'il existe des similarités entre ces différents documents, et si l'on peut constituer une base de données qui relie les considérants des décisions de justice à leurs considérants de principe. Nous verrons que la méthode issue de la comparaison avec le cosinus présente cependant des inconvénients et ne nous permet pas d'avoir une certitude sur les résultats obtenus.

Ce n'est qu'en constituant ce nouveau produit issu de notre corpus initial, que l'on peut s'attacher à décrire les phénomènes qui caractérisent les jurisprudences : à la fois dans leur ensemble, et dans leurs utilisations individuelles par la justice administrative. Ces deux perspectives nous permettront d'avoir une vision macroscopique des phénomènes juridiques qui traversent la justice administrative, ainsi qu'une vision microscopique sur les utilisations des jurisprudences. Nous verrons comment nous pouvons produire des connaissances grâce au traitement automatique du langage, et aux techniques statistiques, et nous présenterons et analyserons une sélection de résultats.

A. Traiter le corpus pour le rendre analysable : vectoriser le texte

1. Rendre le texte comparable

Notre corpus d'un peu plus de 5.000.000 considérants a été constitué pour la comparaison, comme nous l'avons évoqué dans la partie précédente. Mais nous n'avons pas encore transformé les documents en vecteurs qui seront à même d'être traités avec des outils statistiques et mathématiques. Cette étape suppose d'abord une réflexion sur les mots qui doivent être conservés, transformés, écartés, pour permettre un meilleur traitement de ces documents. En effet, la sur-représentation de certains mots dans le corpus peut amener, par exemple, des ressemblances entre deux considérants qui n'ont pas forcément lieu d'être. On peut notamment essayer de constituer des formes « d'entités nommées » pour éviter la sur-représentation de certains termes.

Pour rendre comparable les documents du corpus, nous devons donc de nouveau re-coder en partie les données. En particulier, nous supprimons toute une liste de « mots-outils », qui n'ont pas forcément d'intérêt dans la mise en comparaison, voire qui peuvent biaiser celle-ci. En particulier, nous avons retiré des considérants les mots-outils français que l'on retrouve au sein de la librairie nltk : des pronoms, des pronoms personnels, les déclinaisons des verbes être et avoir, certaines lettres, auquel nous ajoutons les mots « considérant » et « qu », notamment parce qu'on les retrouve pour chaque considérant, et qu'ils n'ont donc pas d'importances, voire tendent à rendre similaires les documents qui ne le seraient pas. Nous transformons par la suite certains mots en tokens⁵⁶, en les ré-écrivant sous une autre forme : c'est le cas des différents codes juridiques, dont nous avons réalisé une liste. C'est également le cas pour les références à des articles, des décrets, des ordonnances, ou des lois, ainsi que pour les dates et les citations. Nous ré-écrivons également quelques abréviations⁵⁷.

Nous avons décidé de « tokenizer » ces éléments, afin de limiter les interprétations des mots les plus courants, notamment les mots codes, et articles. Cela permet de restreindre le sens des mots, mais également de les associer directement à leur objet. De cette manière les signifiants n'ont qu'une seule signification. En ce qui concerne les dates

⁵⁶ C'est-à-dire en créant de nouveaux objets textuels qui ne sont pas des mots.

⁵⁷ Ces opérations sont observables dans les fichiers `jade_rewrite.txt`, `jade_patterns.txt`, `jade-articles.txt`, et `jade_codes.txt`, qui sont parmi les éléments du code source.

nous avons décidé de les re-coder afin de faire en sorte que les nombres soient bien associés à un mois ou à une année, afin de ne pas les enlever des documents, car nous supprimons tous les nombres et chiffres qui ne sont pas associés à un « token ». Dans le cas des articles, nous conservons les chiffres et nombres des alinéas et des chapitres. Nous avons souhaité supprimer les chiffres, car une jurisprudence ne mentionne presque jamais de chiffre ou de nombre, mais explicite une règle de Droit. En ce sens, les chiffres ne sont pas utiles à la comparaison, ou ils ce sont des passages des paragraphes qui ne sont pas des reprises de la jurisprudence. Nous n'avons pas souhaité lemmatiser ou stemmatiser le texte dans la mesure où : « *les structures des textes juridiques sont des points d'appui précieux pour construire une analyse automatique* »⁵⁸, et que perdre une trop grande partie de cette structure linguistique ne nous paraît pas judicieux.

En ce qui concerne les citations, nous avons souhaité les supprimer. Cela peut paraître contre-intuitif, notamment car les lois sont parfois citées au sein des considérants qui les mobilisent, en même temps qu'elle invoque la jurisprudence par ailleurs. Cependant, les citations n'ont pas d'intérêts dans le cadre de ce travail, dans la mesure où les jurisprudences ne sont pas constituées de citations du texte de loi, mais des phrases ou des paragraphes qui apportent un raisonnement sur la loi, et qui l'appliquent dans un certain sens. En ce sens, les citations d'articles, de lois, ou de tout autres documents juridiques ne nous sont pas non plus utiles, puisque l'on cherche les références aux jurisprudences, et que nous conservons les références aux documents juridiques dans le texte d'autre part. Nous ajoutons un traitement supplémentaire pour les considérants de principe des décisions du Conseil d'Etat, car ils contiennent également les codes PCJA. Nous transformons ces derniers en token, afin de pouvoir les identifier et les extraire par la suite.

C'est à partir de ce traitement préalable que nous pouvons maintenant transformer nos documents en vecteurs. Dans notre cas, et notamment parce que les objets qui sont à comparer sont les mots des phrases des différents considérants, nous avons choisi de travailler avec la technique des bag-of-words. Cette technique permet de transformer des phrases en vecteurs de la longueur du « dictionnaire », c'est-à-dire de l'ensemble des mots de notre corpus. Ainsi, les mots des documents permettent de construire des coordonnées de vecteur. Cette approche, très « simpliste » par rapport à d'autres techniques présentent pour nous plusieurs avantages. D'abord, elle prend bien en considération chacun des mots, leurs multiplicités dans un seul vecteur, et n'a pas

⁵⁸ Clément, Marc, *Art. Cit.*

d'autres objets de construction, ce qui n'est pas le cas par exemple des BoW TF-IDF. Par rapport à des embeddings, c'est-à-dire dans notre cas des vecteurs construits à partir de modèles pré-entraînés, les BoW sont plus à même de mieux fonctionner pour le calcul de similarité avec une variété de mots : « *Especially in settings where two or more words, which have distinct meanings, are likely to be mentioned close to each other, word embeddings may fail to adequately distinguish between them* »⁵⁹.

La technique du BoW présente cependant plusieurs défauts pour nos vecteurs : elle ne prend pas en compte l'ordre de mots, et donc leur contexte, ce qui serait particulièrement intéressant dans notre cas, et ce qui serait possiblement réalisable avec des embeddings. Les vecteurs sont particulièrement longs en fonction de la taille du corpus, et nous avons donc des vecteurs à 467060 dimensions, ce qui serait également considérablement réduit avec des embeddings, et supposerait un temps de calcul beaucoup plus court pour la comparaison entre deux vecteurs. Cependant, on prend ici en compte la fréquence et la présence des mots dans les documents afin de les comparer et c'est précisément ce qui nous intéresse étant donné que nous recherchons les réutilisations des jurisprudences dans les décisions de justice : nous cherchons donc pour chaque considérant de principe, les considérants des décisions de justice qui partagent le plus de mots possibles avec lui.

2. Automatiser la recherche des jurisprudences

Pour réaliser cette opération, nous effectuons un calcul de similarité fondé sur le cosinus entre les vecteurs. Plus le cosinus tend vers un, plus les vecteurs sont ressemblants, et à l'inverse, plus le cosinus est proche de zéro, moins les vecteurs se ressemblent. Pour voir comment les jurisprudences sont utilisées au sein de la justice administrative, il nous faut systématiser ce travail de comparaisons. Nous avons donc choisi de créer un tableau qui met en relation les considérants de principe (les jurisprudences), avec les considérants de décisions. Il nous faut d'abord comprendre les

⁵⁹ Barua, Josh. « Word Embeddings Versus Bag-of-Words: The Curious Case of Recommender Systems ». *The Startup*, 25 août 2020, <https://medium.com/swlh/word-embeddings-versus-bag-of-words-the-curious-case-of-recommender-systems-6ac1604d4424>.

difficultés que l'on peut rencontrer dans le calcul de similarité. A partir d'ici, nous parlerons de jurisprudence pour évoquer les considérants de principe⁶⁰.

En fonction de la variété des mots qui sont employés dans deux considérants de décisions de justice, quand bien même ils citeraient la même jurisprudence, ils n'auront pas le même « score » avec cette dernière. La valeur du cosinus est relative en fonction du corpus sur lequel on travaille. Dans notre cas, des documents peuvent être ressemblants à une jurisprudence et n'avoir qu'un cosinus de 0.25, dans la mesure où d'autres mots entourent notre considérant de décision, et que sa syntaxe est plus ou moins différente de la rédaction de la jurisprudence qu'il cite. La valeur du score est donc très relative, il n'existe donc pas de valeur « plancher » du cosinus à partir de laquelle nous sommes sûr de récupérer les considérants qui citent la jurisprudence que nous testons. C'est-à-dire que plus nous abaissons le seuil du cosinus, plus nous récupérons de « bons » considérants et de « mauvais » considérants.

En pratique, pour faire les associations entre les considérants et leurs jurisprudences, nous ne sommes pas obligés de fixer un seuil dans la mesure où nous utilisons une technique qui nous donne une valeur de cosinus pour chaque considérants de notre base pour un considérant de principe donné. Nous regardons donc simplement quel considérant de principe « score » le mieux, et faisons le même travail avec un considérant vers les considérants de principe. Mais si l'on conserve toutes les valeurs du cosinus, le travail de « fouille », prend beaucoup plus de temps, et nous ne pensons pas nécessaire de conserver les valeurs pour lesquelles nous savons très probablement que ce travail ne fonctionne pas, par exemple en dessous de 0,1. Même si nous avons tenté de voir s'il existait un minimum de la valeur du cosinus où nous avons une majorité d'individus qui correspondent à la jurisprudence citée, ces méthodes reposent finalement sur une appréciation arbitraire du « bon » cosinus, qui dépend des documents mis en comparaison. De plus, la longueur du considérant (et donc la présence ou non d'une diversité de mots) influe fortement sur la valeur du cosinus : plus un document est court, plus il a de chance, s'il partage des mots en commun, d'avoir des valeurs importantes du

⁶⁰ Nous parlons ici de jurisprudence pour faire la différence entre les considérants de principe qui ne sont pas ré-utilisés par les juridictions, et ceux qui le sont, qui acquiert, en quelque sorte, le statut de jurisprudence. De plus, cela facilite la distinction pour le lecteur entre les considérants des décisions, et les considérants de principe.

cosinus, alors même que les documents ne partagent par exemple que 3 mots⁶¹. De plus, les documents très courts sont très souvent l'annonce d'une conclusion.

Le travail d'association des jurisprudences avec les considérants des décisions de justice qui les citent nécessitent d'être le plus optimisé possible, dans la mesure où il peut prendre beaucoup de temps à être réalisé. Nous réalisons donc un ensemble de tâches au sein du processus d'association afin de le simplifier, et d'au mieux récupérer les « bonnes » associations, c'est-à-dire, les considérants qui citent effectivement la jurisprudence ou qui s'en rapprochent le plus. Nous regardons pour chaque jurisprudence :

1. Si elle comporte plus de 10 mots après les pré-traitements que nous avons mentionné. Nous estimons que si les considérants de principe ont 10 mots ou moins, ils ne peuvent pas être considérés comme des jurisprudences.
2. Nous conservons les considérants de principe qui ne sont pas des ré-écritures exactes d'une autre, au sein d'un même fichier. Il arrive qu'un considérant de principe soit ré-écrits plusieurs fois dans la même décision, en analyse, sous l'exacte même formulation, avec une différence au niveau du code PCJA. Nous ne prenons donc pas en compte les doublons.
3. Nous conservons les considérants de principe qui ont au moins un considérant avec une valeur supérieure à 0.7 de cosinus avec eux. Nous réalisons cette opération pour ne sélectionner que les considérants de principe qui sont effectivement des jurisprudences, c'est-à-dire qu'elles sont ré-employées par les juridictions.
4. Nous conservons uniquement les considérants dont la valeur du cosinus est supérieure ou égale à 0.27 dans la similarité avec la jurisprudence.
5. Nous enlevons de cet ensemble de considérants tous ceux qui moins de 10 mots après pré-traitement.
6. Nous conservons uniquement les considérants dont le calcul de similarité sur l'ensemble des jurisprudences, ont le cosinus le plus élevé avec la jurisprudence

⁶¹ Voici par exemple deux documents qui partagent 0.48507124 de cosinus :

1. « Au fond : »

2. « 54-08-02-03-02-01 Le juge de cassation peut procéder à une substitution de motifs en écartant un moyen au motif qu'il est inopérant, alors que le juge du fond l'avait rejeté au fond. »

Les deux documents partagent 1 mot, mais le considérant ne citant pas la jurisprudence, on ne peut pas savoir s'ils sont en effet associé. On voit ici que la différence de taille influe sur la valeur du cosinus.

que nous testons. Cela permet de renforcer l'assurance de la similarité entre le considérant et la jurisprudence.

En plus de ces opérations, nous ne prenons pas en compte les jurisprudences où nous retrouvons moins de 6 considérants qui lui sont associés. De cette manière, nous essayons de traiter les jurisprudences qui sont utilisés. Ce travail de nettoyage permet une nouvelle fois de gagner du temps de traitement, ainsi que d'éliminer le plus possible les « pollutions » et les redondances (doublons) pour notre analyse. Au final, sur les 73151 considérants de principe initiaux, nous avons récupéré 11491 jurisprudences, que l'on retrouve dans 1.115.540 considérants de décisions de justice, sur 5.010.748. Ce processus pose néanmoins une question sur les jurisprudences proches textuellement, c'est-à-dire des jurisprudences qui partagent les mêmes mots, ou qui portent sur les mêmes objets : ce sont parfois des évolutions de la jurisprudence dans le temps, ou des ré-écritures avec l'apparition de nouveaux contentieux. La distinction entre deux jurisprudences peut parfois être difficile à réaliser, notamment dans le calcul de similarité. Nous pouvons donc étudier celles-ci ensembles, voir les re-grouper dans les analyses, d'autant plus que nous pouvons dans ce cas observer l'évolution d'une jurisprudence dans le temps. Nous verrons dans quelle mesure cela peut affecter l'analyse dans le B.2 de cette partie. Ce travail d'association permet ainsi de mesurer les utilisations des jurisprudences sur plusieurs variables. Pour montrer les utilisations que l'on peut en faire, nous présentons ici quelques résultats d'analyses.

B. Exemples d'analyses

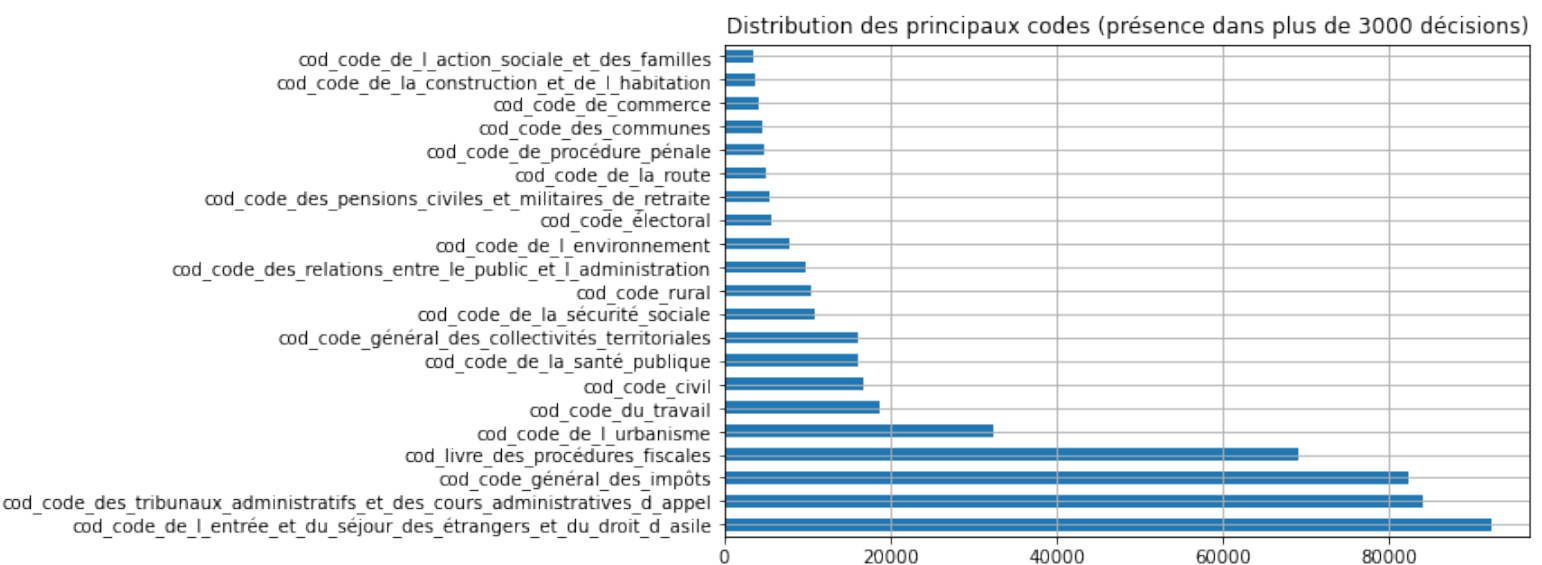
Nous obtenons ainsi un tableau de plus d'un million d'entrées, avec l'identifiant du considérant issu d'une décision, la jurisprudence à laquelle il se rapporte, ainsi que le nom du document où l'on peut retrouver le considérant. De cette manière, on peut réaliser des jointures avec nos autres métadonnées, c'est-à-dire les regrouper dans un seul tableau. De cette manière, on peut réaliser plusieurs tableaux qui regroupent les métadonnées en fonction du considérant ou de la jurisprudence : la date de la décision, la juridiction, le type de recours, le fichage au recueil Lebon, le code PCJA associé à la jurisprudence, ainsi que ses différentes « échelles », les références légales retrouvées dans les visas de la décision. Pour cette partie, nous ne présentons pas toutes les analyses que nous avons pu réaliser dans le cadre du stage, mais nous prendrons

l'exemple de 2 jurisprudences pour illustrer notre propos, et donner à voir l'étendue des connaissances que l'on peut produire, de manière non exhaustive.

1. Étude sur la concentration des considérants sur quelques jurisprudences.

Dans un premier temps, nous pouvons observer des tendances simples à analyser des jurisprudences, avant de s'intéresser à des phénomènes en particulier. En groupant les résultats du tableau des associations par les identifiants des jurisprudences, on peut chercher à voir quelles jurisprudences sont les plus utilisées⁶². On voit que 7 jurisprudences sont mentionnées par plus de 10.000 considérants chacun, sur les 11491 jurisprudences de notre tableau. 10022 jurisprudences ont moins de 100 considérants associés, 1189 ont entre 100 et 500 considérants associés et 280 ont 500 ou plus de considérants qui leur sont associés. On voit que d'après nos travaux, une très petite partie des jurisprudences sont utilisées largement par les juridictions. Une question se pose face à ce constat : est-ce que cette concentration des jurisprudences est liée à l'importance de certains contentieux ?

On peut répondre à la première question en s'intéressant plus particulièrement aux principales jurisprudences, aux codes PCJA⁶³, et aux références légales que l'on retrouve dans les décisions de justices où la jurisprudence est citée. Tout d'abord, on peut regarder si certaines références légales reviennent plus d'autres, ce qui pourrait nous donner une idée des sujets qui sont les plus traités.



⁶² Voir annexe 1.8 : Graphique des jurisprudences les plus utilisées.

⁶³ Plan de Classification de la Justice Administrative

Ce graphique nous montre que quelques codes sont particulièrement utilisés au sein de la justice administrative : le code de l'entrée et du séjour des étrangers et du Droit d'asile, le code des tribunaux administratifs et des cours administratives d'appel, le code général des impôts, le code de procédures fiscales et dans une moindre mesure le code de l'urbanisme. On peut écarter le code des tribunaux administratifs qui ne nous donne pas particulièrement d'information quant aux contentieux qui sont les plus souvent traités par la justice administrative. En revanche on constate que deux thématiques ressortent particulièrement : les affaires fiscales et le contentieux des étrangers. Pour voir si l'on retrouve ce constat dans les jurisprudences, on peut regarder sur quel(s) sujet(s) portent les 7 principales jurisprudences, et avec quels codes ils reviennent le plus souvent⁶⁴. On observe ainsi que 3 jurisprudences concernent le contentieux des étrangers, et les 4 autres concernent les frais et dépenses liés au procès, que l'on retrouve dans le code de justice administrative. On le supprime également des résultats pour dégager des tendances sur les contentieux (et non les mesures de procédures) pour la suite de ce travail. Pour obtenir des résultats sur les codes, nous avons réalisé une jointure entre l'identifiant de la décision des considérants associés à une jurisprudence, et les références légales de cette décision. De cette manière, nous avons les différents codes qui sont mobilisés pour chacun des considérants, que l'on rapporte à la jurisprudence qu'ils citent. Nous avons donc des résultats qui ne représentent pas concrètement la réalité, mais qui nous donnent des tendances sur l'utilisation des jurisprudences.

On peut regarder si les quatre jurisprudences qui concernent les frais se retrouvent plus souvent sur un ou plusieurs codes en particulier pour voir si ces jurisprudences sont exclusives ou non. Prenons l'exemple de la première, qui porte l'identifiant 35553⁶⁵ : On constate d'abord qu'en termes de pourcentage, il n'y a pas de tendances particulières. La jurisprudence ne s'impose pas avec un code en particulier, et on la retrouve dans un peu moins de 1% des considérants qui sont dans une décision qui mentionne le code la recherche, et dans un peu plus de 14% des considérants qui sont associés, dans leurs décisions, au code du domaine de l'Etat. Cela signifie, que 14% des considérants qui sont reliés au code du domaine de l'Etat citent cette jurisprudence, ce qui est loin d'être négligeable, même si ce code ne représente que 702 considérants. Pour relativiser ces résultats, on peut regarder la répartition des codes associés aux considérants affiliés à la

⁶⁴ Voir annexe 2.2 : Exemples des 7 jurisprudences les plus mentionnées

⁶⁵ Voir annexe 1.8 : Graphiques des références légales de la jurisprudence 35553

jurisprudence 35553. On remarque alors que le code du domaine de l'Etat représente une infime partie des codes avec lesquels est utilisé cette jurisprudence, à l'inverse des quatre codes les plus utilisés. On retrouve des résultats à peu près similaires pour les autres jurisprudences qui s'intéressent au frais et aux dépenses liés au procès et au litige.

En revanche on constate un phénomène différent pour les jurisprudences qui portent sur le contentieux des étrangers. Si l'on s'intéresse à la jurisprudence 32771⁶⁶, qui concerne le Droit au respect de la vie familiale, dans le cadre de l'article 8 de la Convention européenne de sauvegarde des Droits de l'homme et des libertés fondamentales, qui permet à un individu d'annuler la demande d'expulsion dont il fait l'objet⁶⁷, on remarque que les pourcentages, tout comme la répartition des références aux codes, sont très largement dominés par le code de l'entrée et du séjour des étrangers et du Droit d'asile. 8% des saisies de ce code se font avec cette jurisprudence, ce qui représente plus de 30000 considérants. Cela signifie que dans notre corpus, un peu moins d'une affaire sur 10 qui concerne le code de l'entrée et du séjour des étrangers et du Droit d'asile invoque cette jurisprudence.

Pour étayer ces analyses et répondre à notre interrogation, on peut également s'intéresser aux codes PCJA, qui recensent les différents contentieux auxquels est confronté la justice administrative. Si nous n'avons pas le code PCJA des considérants des décisions, nous avons les codes PCJA pour les jurisprudences, et nous avons l'association de celles-ci avec les considérants. Nous pouvons donc associer le code PCJA d'une jurisprudence aux considérants qui l'invoque. Même si cette nouvelle association risque un biais cognitif⁶⁸, nous considérons que la mention de la jurisprudence suppose un cas d'application semblable, ou au moins très proche. Cette opération suppose une autre difficulté : les jurisprudences peuvent avoir plusieurs codes PCJA, puisqu'elles peuvent porter sur plusieurs contentieux. Cela suppose que nous

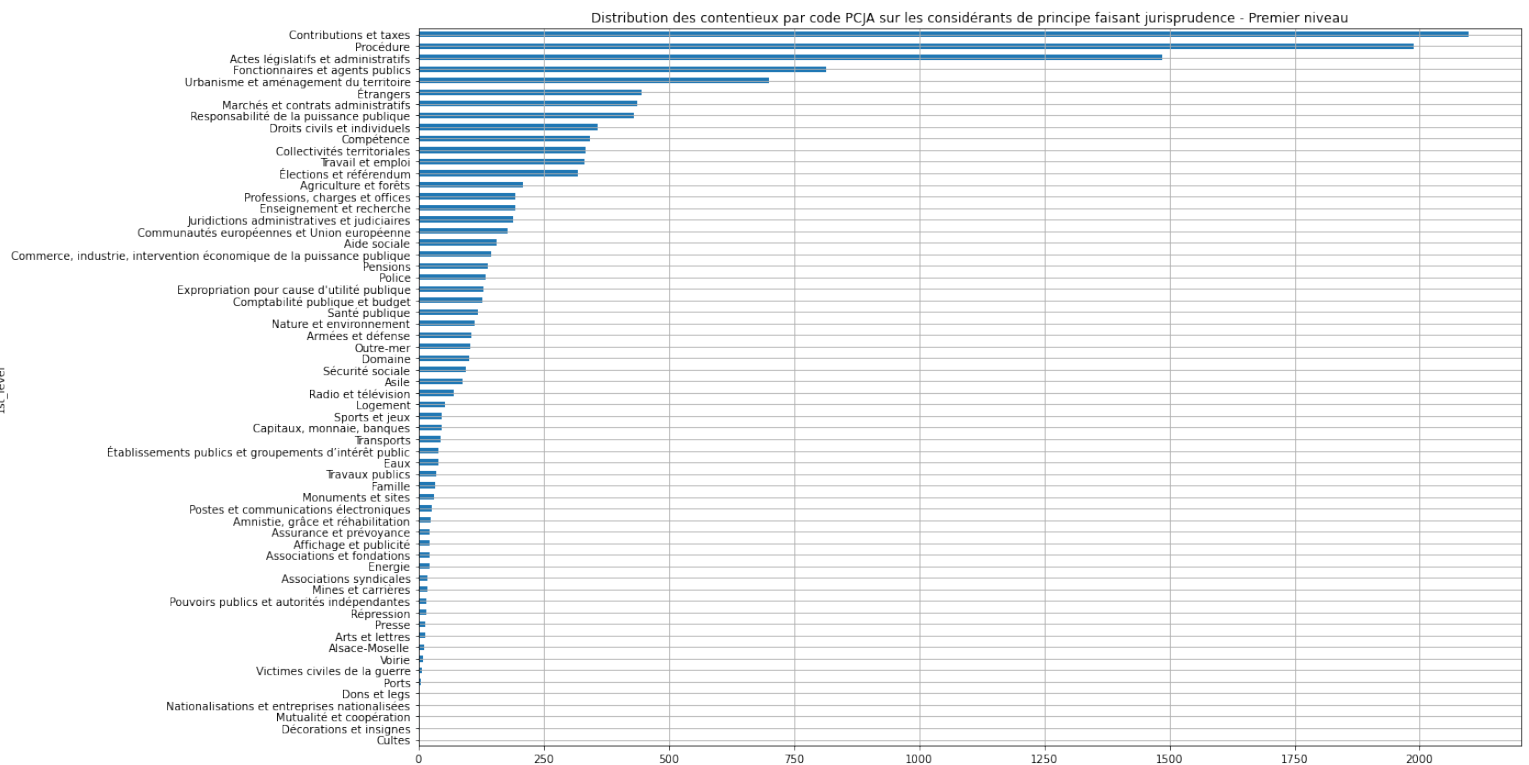
⁶⁶ Voir annexe 1.9 : Graphiques des références légales de la jurisprudence 32771

⁶⁷ « Un étranger peut utilement se prévaloir, à l'appui de conclusions tendant à l'annulation de la mesure d'expulsion dont il a fait l'objet, des stipulations de l'article 8 de la Convention européenne de sauvegarde des Droits de l'homme et des libertés fondamentales aux termes desquelles : "1°) Toute personne a Droit au respect de sa vie privée et familiale, de son domicile et de sa correspondance - 2°) Il ne peut y avoir ingérence d'une autorité publique dans l'exercice de ce Droit que pour autant que cette ingérence est prévue par la loi et qu'elle constitue une mesure qui, dans une société démocratique, est nécessaire à la sécurité nationale, à la sûreté publique, au bien-être économique du pays, à la défense de l'ordre et à la prévention des infractions pénales, à la protection de la santé ou de la morale, ou à la protection des Droits et libertés d'autrui" ».

⁶⁸ Cela suppose que l'application de la jurisprudence porte toujours sur le ou les mêmes contentieux.

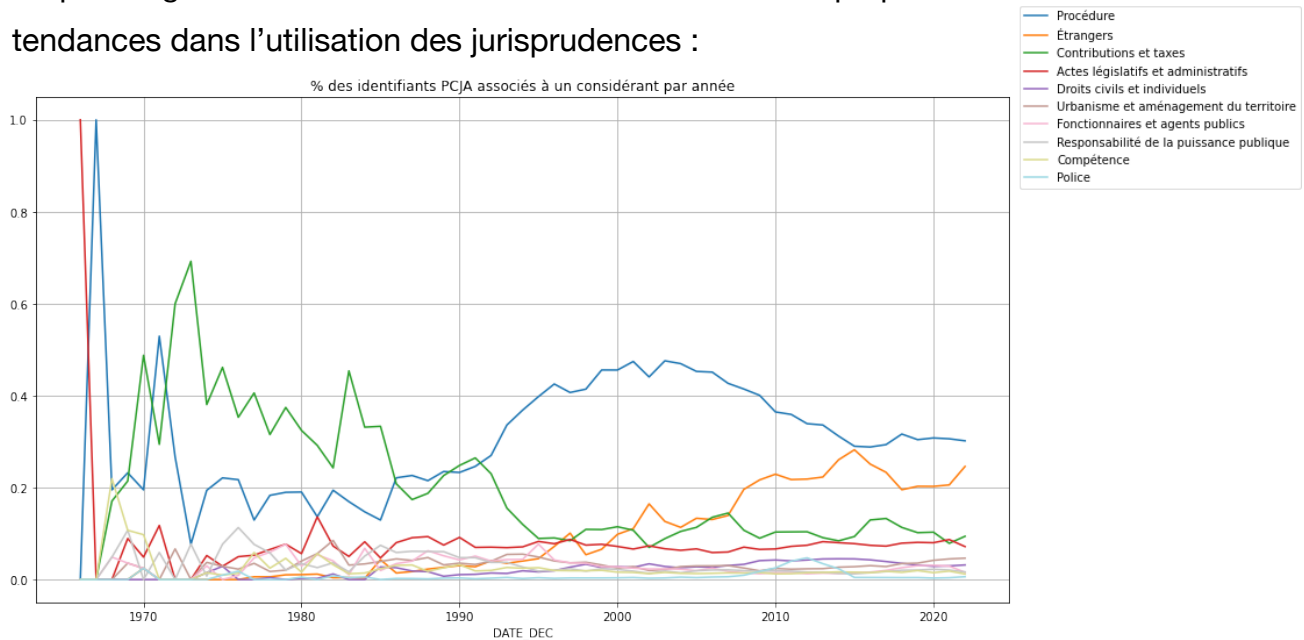
devons soit faire un choix dans l'attribution d'un code PCJA au considérant, soit que l'on duplique les considérants pour représenter l'importance des codes PCJA. Nous avons opté pour cette dernière option, dans la mesure où il nous était impossible de discriminer l'attribution d'un code PCJA, à moins de lire un par un chaque considérant associé à une jurisprudence. Ainsi, nous pouvons décrire des tendances au sein des contentieux de la justice administrative, qui restent relative aux jurisprudences.

Nous avons donc deux axes possibles de lecture des contentieux : ceux qui ont le plus de jurisprudence, et ceux qui sont, a priori, les plus utilisés, c'est-à-dire qui ont le plus de considérants associés. Sur la première ligne d'analyse, on peut se questionner sur les contentieux qui ont le plus de jurisprudences, ce qui peut nous renseigner sur les contentieux qui évoluent le plus. On peut d'abord regarder le « 1er niveau » du code PCJA, afin de voir les principales thématiques des jurisprudences.



On remarque ainsi que c'est le contentieux sur les contributions et taxes, ou contentieux fiscal, suivi de près par le contentieux qui concerne la procédure, et d'un peu plus loin par celui qui concerne les actes législatifs et administratifs sont les contentieux qui ont le plus de jurisprudences. Ici, le contentieux des étrangers, bien qu'ils soient assez importants, n'est pas au premier plan, c'est-à-dire qu'il y a moins de 500 jurisprudences qui concernent ce contentieux. Il nous faut maintenant voir les contentieux les plus

on peut regarder l'évolution des contentieux dans le temps pour voir si l'on retrouve des tendances dans l'utilisation des jurisprudences :



Ce dernier graphique renforce nos analyses précédentes : on remarque que certains contentieux sont particulièrement utilisés, et donc mobilisateurs de certaines jurisprudences. En particulier le contentieux des étrangers, qui devient de plus en plus important avec le temps, où encore certaines jurisprudences qui concernent la procédure. Le contentieux fiscal, qui étaient très important entre les années 1970 et 2000, devient un sujet de de moins en moins important, en proportion, dans le temps. On peut également constater que si les proportions diminuent pour la plupart des contentieux « historiques », c'est également parce que la diversité des contentieux augmente avec le temps, comme on peut le remarquer avec le graphique.

Ainsi, le fait que certains contentieux soient plus récurrents que d'autres, amènent la forte ré-utilisation de certaines jurisprudences qui portent sur ces contentieux. On retrouve ces tendances dans les rapports d'activités de la justice administrative : Au sein des juridictions du fond et des cours d'appels, le contentieux des étrangers est particulièrement important, suivi par les contributions et les taxes. De pouvoir voir quelles jurisprudences sont particulièrement utilisées dans ces contentieux permet ainsi

d'objectiver des savoirs plus ou moins connus et reconnus par les juges et la justice administrative dans son ensemble.

2. Étude d'un cas particulier de jurisprudence : limites du travail d'analyse

En plus de ces informations générales issues de l'analyse de l'ensemble des jurisprudences, nous pouvons également chercher à analyser une jurisprudence en particulier, ou un ensemble de jurisprudences « proches ». Dans notre cas, nous souhaitons étudier une jurisprudence qui concerne le licenciement des salarié.es investi.es d'une fonction représentative, qui bénéficient d'une protection spécifique contre le licenciement dû à cette fonction de représentation. Nous prenons donc en compte la jurisprudence qui porte l'identifiant 49511. Cette jurisprudence et les jurisprudences qui lui sont « proches » par rapport à la similarité cosinus, représentent 2815 considérants issus des décisions de justice. On peut d'abord chercher à voir à quoi ressemblent les différentes jurisprudences qui lui sont proches, en quelles années elles ont été créées, et combien de considérants elles représentent.

On remarque tout d'abord que pour une même situation de Droit (la qualité exceptionnelle des travailleurs exerçant une fonction représentative), nous avons différents **cas d'espèce** : c'est-à-dire différentes situations pour lesquelles le Conseil d'État a dû s'exprimer pour trancher un litige. En l'occurrence, la jurisprudence 49511 s'intéresse au comportement du salarié en dehors du contrat de travail :

« Dans le cas où la demande de licenciement est motivée par un acte ou un comportement du salarié survenu en dehors de l'exécution de son contrat de travail, notamment dans le cadre de l'exercice de ses fonctions représentatives, il appartient à l'inspecteur du travail, et le cas échéant au ministre, de rechercher, sous le contrôle du juge de l'excès de pouvoir, si les faits en cause sont établis et de nature, compte tenu de leur répercussion sur le fonctionnement de l'entreprise, à rendre impossible le maintien du salarié dans l'entreprise, eu égard à la nature des fonctions de l'intéressé et à l'ensemble des règles applicables à son contrat de travail. ».

Ce n'est pas le même cas pour la jurisprudence 64008 qui cite même le cas d'espèce sur lequel elle tranche :

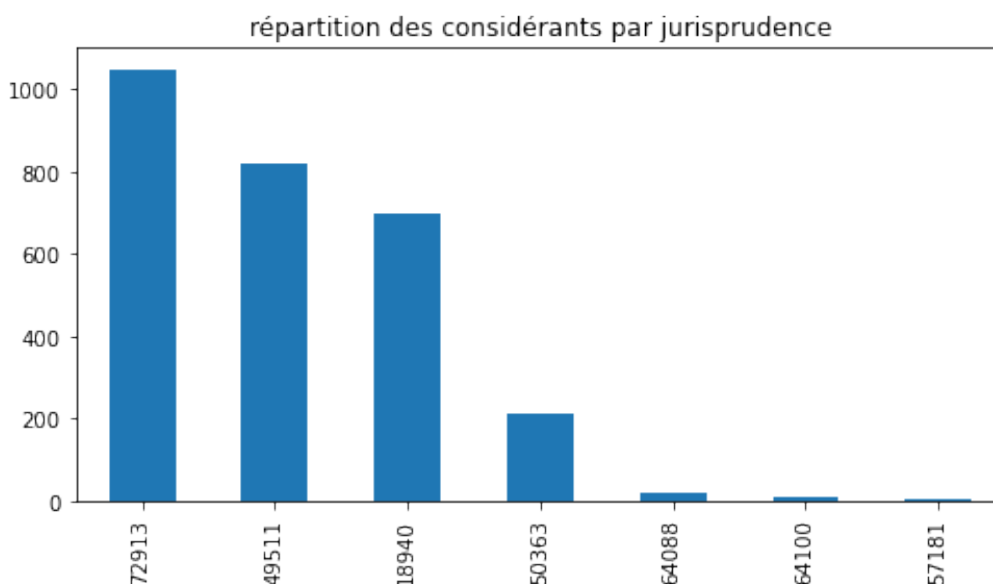
« Dans le cas où la demande de licenciement est motivée par un comportement fautif, il appartient à l'inspecteur du travail saisi et, le cas échéant, au ministre compétent, de rechercher, sous le contrôle du juge de l'excès de pouvoir, si les faits reprochés au salarié sont d'une gravité suffisante pour justifier le licenciement, compte tenu de l'ensemble des règles applicables au contrat de travail de l'intéressé et des exigences propres à l'exécution normale du mandat dont il est investi. Un agissement du salarié intervenu en-dehors de l'exécution de son contrat de travail ne peut motiver un licenciement pour faute, sauf s'il traduit la méconnaissance par l'intéressé d'une obligation découlant de ce contrat.,,2) En l'espèce, salarié protégé ayant, lors d'une suspension de séance du comité d'établissement, assené un violent coup de tête à un autre salarié. Cet acte de violence délibérément commis sur la personne d'un collègue sur le lieu du travail, même à l'occasion des fonctions représentatives de l'intéressé, doit être regardé comme une méconnaissance par celui-ci de son obligation, découlant de son contrat de travail, de ne pas porter atteinte, dans l'enceinte de l'entreprise, à la sécurité d'autres membres du personnel. »

On remarque donc que chacune des jurisprudences sont bien liées à la question du licenciement des travailleurs exerçant des fonctions représentatives, avec cependant des applications à des cas différents : comportement fautif, comportement en dehors de l'exécution du contrat de travail, et dans le cadre des fonctions représentatives. On constate que dans ce groupe de jurisprudence, ce sont des évolutions ou des spécifications de l'application de la règle de Droit que l'on observe. Ce constat est renforcé quand on s'intéresse aux dates des jurisprudences, qui suivent une chronologie : de 1976 à 2021, en s'arrêtant en 2005, en 2010 et en 2015. Cela renforce l'idée d'une progression, d'une adaptation, de l'application du Droit sur ce sujet.

Quand on s'intéresse à l'évolution de l'utilisation de ces jurisprudences dans le temps⁷⁰, on remarque, sans grande surprise, que le Conseil d'État est saisi de ce sujet jusqu'en 1999, avant que les cours d'appels ne prennent le relais. Cependant, quand on observe l'utilisation des différentes jurisprudences par les différentes juridictions⁷¹, on remarque que le Conseil d'État a tendance à utiliser plus souvent la jurisprudence la plus ancienne, et que c'est la jurisprudence la plus récente (72913) qui est la plus utilisée au total :

⁷⁰ Voir annexe 1.11 : Distribution des considérants associés à la jurisprudence 49511, et jurisprudence proches, en fonction de la juridiction

⁷¹ Voir annexe 1.12 : Distribution de l'ensemble des jurisprudences proche de 49511, en fonction de l'année et de la juridiction



On peut s'étonner de cette dernière observation, étant donné que cette jurisprudence n'a qu'un an, et qu'il y a peu de chance que celle-ci ait été utilisée plus de fois que les autres, qui sont plus anciennes. Nous sommes d'autant plus étonnés que l'on remarque dans le détails de chaque jurisprudence, que la plus récente est associée à des considérants qui sont dans des décisions de justice de 1986⁷², et que le Conseil d'État aurait prononcé son avis sur ces décisions, c'est-à-dire qu'il y aurait déjà dû y avoir un considérant de principe énoncé avant 2021.

Pour contrôler cette particularité, on peut regarder directement les considérants qui sont associés à la jurisprudence. Elle porte en particulier sur la légalité des procès-verbaux rédigé par un huissier de justice. Quand on regarde les considérants le plus ressemblants et le moins ressemblants, on obtient ces résultats :

« 2. **En vertu des dispositions du code du travail, les salariés légalement investis de fonctions représentatives qui bénéficient, dans l'intérêt des travailleurs qu'ils représentent, d'une protection exceptionnelle, ne peuvent être licenciés qu'avec l'autorisation de l'inspecteur du travail. Lorsque le licenciement d'un de ces salariés est envisagé, il ne doit pas être en rapport avec les fonctions représentatives normalement exercées par l'intéressé ou avec son appartenance syndicale.** Dans le cas où la demande est motivée par un comportement fautif, il appartient à l'inspecteur du travail saisi et, le cas échéant, au ministre compétent, de rechercher, sous le contrôle du juge de l'excès de pouvoir, si les faits reprochés au salarié sont d'une gravité suffisante pour justifier son licenciement, compte tenu de l'ensemble des règles applicables à son contrat de travail et des exigences propres à l'exécution normale du mandat dont il est investi. ». Considérant le plus ressemblant

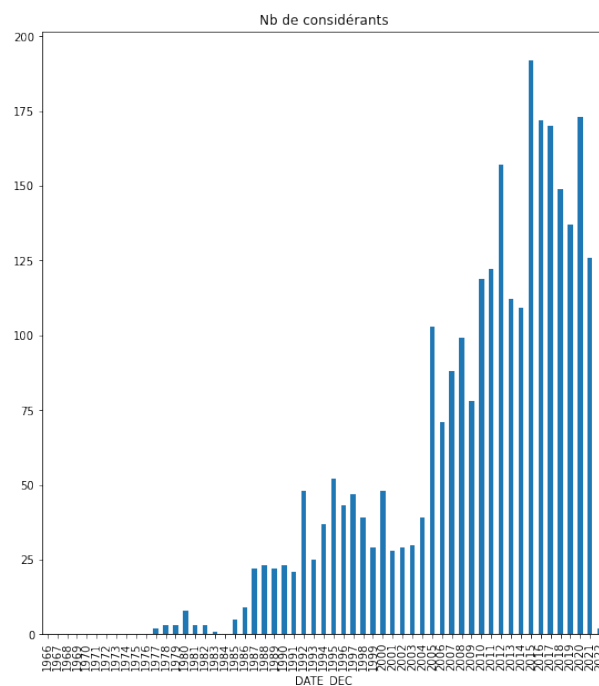
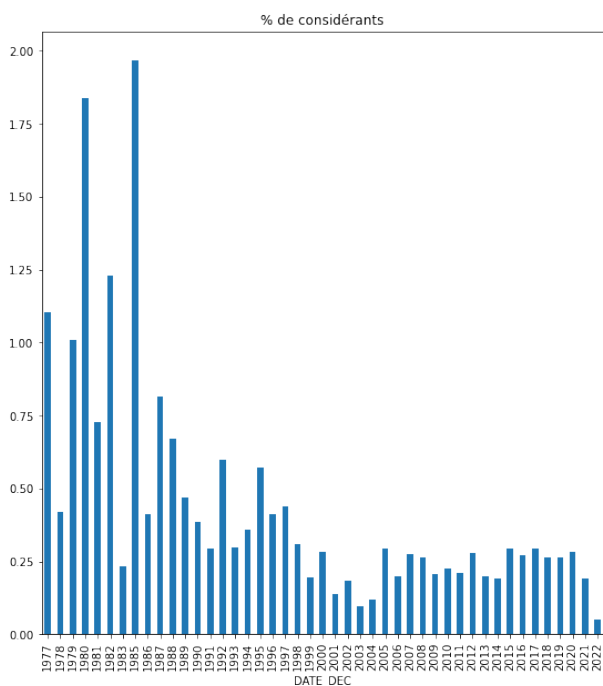
⁷² Voir annexe 1.13 : Distribution des considérants par année et par juridiction pour la jurisprudence 72913 en pourcentage par année

« Considérant qu'en vertu de l'article L. 425-1 du code du travail, les délégués du personnel titulaires et suppléants, qui bénéficient, dans l'intérêt des travailleurs qu'ils représentent, d'une protection exceptionnelle, ne peuvent être licenciés qu'avec l'autorisation de l'inspecteur du travail ; que, lorsque le licenciement d'un de ces salariés est envisagé, il ne doit pas être en rapport avec les fonctions représentatives normalement exercées par l'intéressé ou avec son appartenance syndicale ». Considérant le moins ressemblant

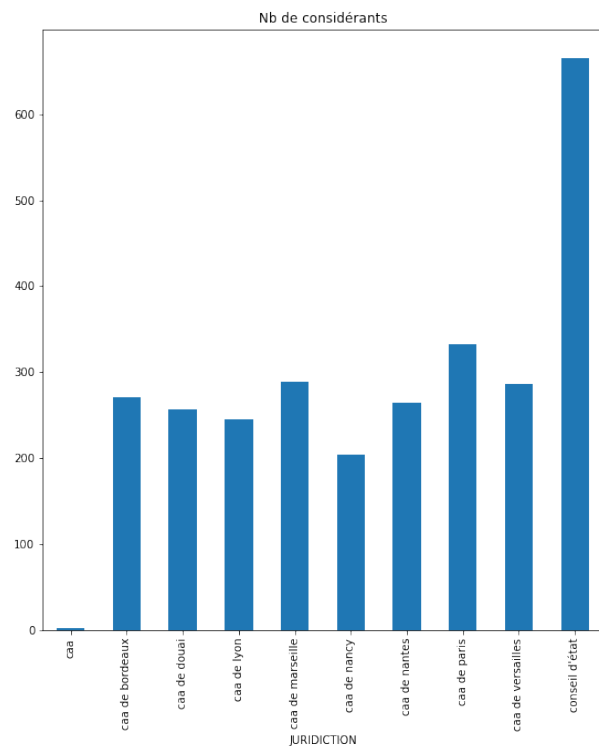
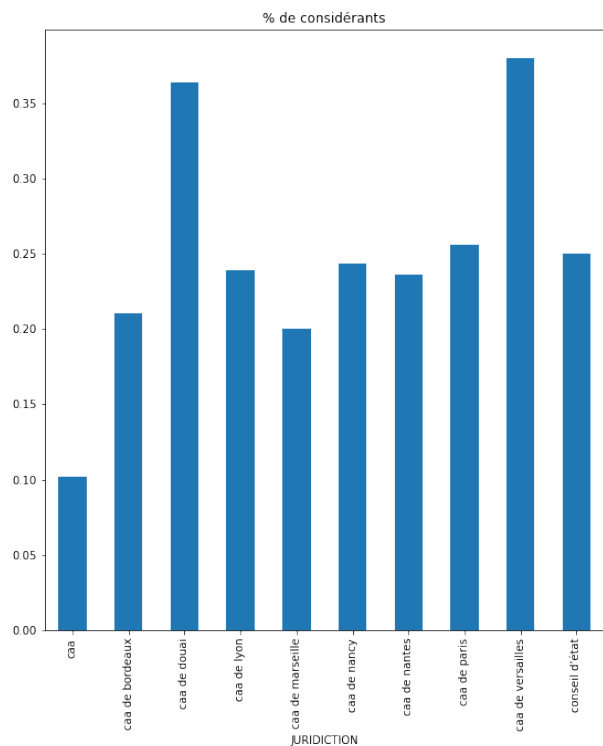
Dans les deux cas, le cas d'espèce n'est pas mentionné, il n'y a pas de mention de la question du procès-verbal de la part d'un huissier, et l'on retrouve la partie commune à ce groupe de jurisprudence sur la procédure de licenciement d'un.e salarié.e exerçant des fonctions représentatives.

On ne peut donc pas affirmer que ces considérants sont effectivement liés à cette jurisprudence. Ils le sont cependant au groupe de jurisprudence. Cette observation est notamment liée à notre manière de constituer, de préparer et d'exploiter notre corpus. En effet, la similarité cosinus fonctionne sur le nombre de mots en commun entre deux documents, et notre manière de découper les considérants de manière automatisée ne nous permet pas de récupérer toutes les informations nécessaires, dans le cas présent, à affiner la connaissance que l'on peut produire sur les jurisprudence au cas par cas. On pourrait remédier à ce problème en ne récupérant pas uniquement les considérants individuellement, mais en sélectionnant les motifs en entier, ce qui donnerait plus de contexte et de mots pour comparer les documents entre eux.

Ceci étant dit, nous pouvons tout de même inférer des comportements sur le groupe de jurisprudence en lui-même. Ou, nous pouvons effectuer d'autres travaux sur des groupes où la distinction entre les jurisprudences est plus marquée. Toujours pour le groupe de considérant proche de la jurisprudence 49511, on peut chercher à voir comment ce dernier évolue au cours du temps :



On remarque que si le nombre de considérants liés à cette jurisprudence augmente au cours du temps, le pourcentage se stabilise rapidement entre 0.25% et 0.5% des considérants par année, même si cela représente autour de 2% des considérants (1 sur 50) en 1985, alors même que le nombre de considérants est très inférieur à 25. En ce qui concerne les juridictions, on peut réaliser les mêmes opérations :



On peut ainsi observer si une juridiction est plus à même de mobiliser, en proportion, un groupe de jurisprudence par rapport aux autres. Ici, on remarque que les cours d'appel de Versailles et de Douai ont plus souvent à juger sur les questions de licenciements des salariés occupant une fonction représentative, par rapport aux autres questions qu'elles traitent. On voit également qu'en cumulé, c'est le Conseil d'État qui a le plus traité ces affaires. On peut relativiser cette dernière affirmation en regardant l'évolution de la répartition des considérants associés à ce groupe de jurisprudence dans le temps et en fonction des juridictions⁷³. Le fait que les cours d'appels n'existaient pas avant 1989 impacte les résultats ci-dessus, et on constate qu'aujourd'hui, la grande partie des considérants proviennent des cours d'appels.

Ces deux exemples tirés de notre travail de stage représentent bien les questions et les enjeux auxquels nous avons été confrontés. Ces questions rejoignent des problématiques d'humanités numériques, comme la transformation des informations en données, ou le sens donné aux analyses et aux visualisations de l'information. Il s'agit en particulier ici de montrer qu'il existe des limites propres aux informations à disposition et aux traitements opérés sur celles-ci. Ces limites n'empêchent pas des analyses poussées sur l'utilisation des jurisprudences présentes dans notre corpus, et ce travail donne une première idée des exploitations possibles des décisions de justice. En ce sens il nous faut interroger ces possibilités au regard du Droit, des évolutions de la recherche et des enjeux inhérents à l'utilisation des données juridiques.

C. Une révolution (épistémologique) du Droit ?

Les analyses que l'on peut tirer de l'étude et de la manipulation des décisions de justices nous amènent à des résultats qui soit pour certains sont déjà connus, voire qui sont déjà produits par le Conseil d'État par d'autres moyens, soit qui semblent nouveaux par rapport aux connaissances que l'on produit ordinairement sur les jurisprudences. On peut douter de l'intérêt de ces analyses, notamment en se demandant s'il est utile de produire des connaissances temporelles du contentieux, alors même que le Conseil d'État peut obtenir ces informations par d'autres biais, et mesure chaque année le volume des contentieux traités. D'un autre côté, peut-on réellement parler de nouveauté dans l'apport de connaissance sur les jurisprudences ? Sur le fonctionnement de la justice

⁷³ Annexe 1.12 : Distribution de l'ensemble des jurisprudences proche de 49511, en fonction de l'année et de la juridiction

administrative ? Ces deux différentes approches méritent une attention particulière : elles permettent de montrer les intérêts et les limites que portent l'analyse, la collecte, et le traitement des données, en sciences humaines et sociales, et de manière générale dans les questionnements de recherche.

1. De « nouvelles » connaissances ?

Sur la question des analyses déjà produites par le Conseil d'État, par d'autres moyens, il est particulièrement intéressant de voir comment les résultats ont été obtenus dans notre cas. En effet, si le Conseil d'État et les tribunaux reposent sur la production d'indicateurs statistiques par dossier pour établir le nombre de décisions rattacher à chaque catégorie de contentieux, nous avons opéré une reconstruction de l'information à partir de l'association entre une jurisprudence (en particulier son ou ses code.s PCJA) et les considérants qui la citent. Ces deux biais différents pour interroger la même information, amènent des tendances sensiblement ressemblantes, mais posent des questions radicalement différentes dans l'obtention des résultats.

En particulier, c'est la question de l'origine de l'information qui pose ici des questions : si elle est relativement simple au sein des analyses produites par le Conseil d'État, elle est beaucoup plus compliquée dans le cas de l'attribution d'un code PCJA à un considérant sur la base de l'affiliation de ce dernier à une jurisprudence. La gymnastique mentale et statistique pour obtenir ces résultats demande un effort de représentation et de compréhension à la fois des outils déployés et de leur potentiel (BoW, similarité cosinus, jointure sur les considérants,...) et une compréhension de l'objet étudié : on comprend alors que ces résultats sont justifiés par le principe même de l'utilisation de la jurisprudence ⁷⁴. Il y a donc en ce sens un vrai travail de questionnement sur l'utilisation des données, et sur l'origine, le sens de celles-ci, comme le montre Bruno Bachimont : « *Il s'agit donc de viser, via les synergies engagées au-delà des disciplines, à mutualiser « les savoirs parcellaires [...] pour former une configuration répondant à nos attentes, à nos besoins et à nos interrogations cognitives (Morin, 1994 : en ligne). Par exemple, l'apparition de ces nouveaux observables et de ces nouvelles pratiques doit conduire à renouveler l'appréhension du sens tel qu'il se constitue dans les discours,*

⁷⁴ Citer la jurisprudence, c'est reproduire une décision sur un procès dont les faits sont ressemblants. On peut donc imaginer que les considérants qui citent la jurisprudence portent sur le même contentieux.

numériques notamment »⁷⁵. C'est par cette « hybridation », ou pluridisciplinarité, que l'on peut parler de nouveauté dans l'analyse, et de renouveau dans la connaissance : on ne produit pas le même savoir avec des ressources différentes. Les données « changent la donne », et permettent de valider ou d'invalider des connaissances, de proposer de nouvelles représentation de celles-ci, ou encore d'explorer des questionnements et des terrains que l'on ne pouvait pas interroger autrement, c'est-à-dire des terrains proprement numériques : « nous avons introduit le calcul depuis longtemps dans nos pratiques scientifiques et culturelles sans qu'on parle de rupture ; c'est depuis que l'on numérise nos contenus culturels, nos outils, nos processus de manière intégrée que nous sommes désormais dans la situation d'aborder n'importe quel objet en mobilisant du calcul sans nous en rendre compte, le calcul étant désormais caché dans les artefacts que nous manipulons. Pour le dire autrement, le numérique apparaît quand d'autres choses que les calculatrices et les ordinateurs (dont la finalité est le calcul) mobilisent le numérique pour d'autres finalités que le calcul. »⁷⁶.

C'est en particulier le cas de l'évolution des jurisprudences dans le temps et l'espace. À l'inverse de la tradition anglo-saxonne du Droit, les jurisprudences ne sont pas explicitement citées dans les décisions de justice, du fait de la question de l'indépendance des juges, et donc du fait que la jurisprudence ne prime pas sur la décision. Dans la pratique, les jurisprudences sont donc citées de manière plus ou moins proche de la rédaction du Conseil d'État. Il n'y a pas de balisage spécifique pour les repérer, de mentions à celles-ci dans les fichiers, et leur repérage peut être particulièrement difficile pour des profanes du Droit. L'analyse dans le temps des jurisprudences dans leur ensemble et dans leurs utilisations individuelles devient alors un travail phénoménal d'annotation, de prospection et de lecture pour des juristes. On comprend dès lors pourquoi l'analyse du contentieux administratif est un travail qui est laissé de côté au sein du Droit, qui privilégie les études qui portent uniquement sur la réalisation de la jurisprudence, et non les usages qui en sont fait par les juridictions.

C'est donc par l'utilisation d'outils de traitement automatique du langage naturel que l'on peut essayer de repérer ces formes particulières du discours et de l'écrit juridique. On voit de nouveau l'intérêt d'une approche pluridisciplinaire qui permet

⁷⁵ Longhi, Julien, *Art. Cit.*

⁷⁶ Bachimont, Bruno. « Le numérique comme milieu : enjeux épistémologiques et phénoménologiques. : Principes pour une science des données ». *Interfaces numériques*, vol. 4, n° 3, 3, décembre 2017, p. 402-402. www.unilim.fr, <https://doi.org/10.25965/interfaces-numeriques.386>.

d'implémenter des solutions informatiques avec une réflexion propre au champ étudié. Cette approche permet ainsi de créer de nouveau terrain permis par le traitement de la donnée : « *Les progrès techniques permettent de proposer au regard de l'esprit des éléments qui seraient sinon dispersés et d'y reconnaître des configurations de sens indécélables sans ces médiations techniques. Les techniques numériques, les « big data », reposent sur ce principe : en rapprochant l'hétérogène, ce dont l'origine, la conformation technique contribuait à séparer et à maintenir dans des univers différents, se retrouve désormais juxtaposé et manipulé dans un même espace pour une synthèse à la fois spatiale (le même espace de stockage) et calculatoire (la mobilisation dans de mêmes calculs)*»⁷⁷. Dans notre cas, l'étude des jurisprudences dans le temps et l'espace, est rendu possible par la mise en place de tel outils, qui permettent ainsi d'interroger des aspects du Droit qui ne pourrait pas l'être autrement. Outils qui présentent également leurs limites, dans leur conception, liées aux contraintes qu'imposent une sélection, et une exploitation des données.

En effet, l'analyse, du fait qu'elle soit construite entièrement sur les corpus de données constitués, est conditionnée par cette construction. Ce constat renforce la nécessité de la pluridisciplinarité d'un travail sur les données, parce qu'elle est la seule manière de contrôler les sens que vont véhiculer les objets manipulés et les limites des travaux. Il s'agit en particulier de trouver les frontières des manipulations pour produire à la fois des connaissances sans pour autant les déterminer. C'est bien toute la complexité du travail sur les données, qui sont à la fois le produit d'une construction arbitraire (au sens des choix qui sont opérés), et d'une reconstruction de sens à l'issue des traitements auxquelles elles sont soumises.

Ces nouvelles approches posent nécessairement la question de la manière de « faire la science ». La transformation des outils, des ressources et des potentiels d'analyses nous permet de revoir nos manières de faire, ou au moins d'élargir les possibilités d'enquêtes. Comment les études sur le Droit peuvent-elles s'adapter à ces évolutions épistémiques ?

⁷⁷ Bachimont, Bruno, « le numérique comme milieu », *Art. Cit.*

2. Faire évoluer l'étude des jurisprudences et des contentieux

En transformant les potentiels de recherche, on peut assister à une réelle révolution épistémologique du Droit : les nouveaux objets d'études, les nouveaux terrains, et les nouveaux outils pour étudier ces derniers amènent à repenser les manières traditionnelles de faire de la recherche en Droit. Peut-on alors encore étudier une jurisprudence, où la création d'une nouvelle jurisprudence, sans prendre en compte les usages qui en sont fait ? C'est ainsi que les possibilités d'analyse du contentieux, que nous avons effleuré dans ce mémoire, viennent transformer le rapport au Droit : on ne fait plus de la recherche en Droit uniquement sur la Doctrine, mais aussi sur les usages du Droit, à savoir sur le contentieux.

Si Bruno Bachimont invite à re-considérer toutes les recherches fondées sur un travail des données en science de la données⁷⁸, les humanités numériques, en prenant en compte les données dans leur contexte de production, amènent au contraire à donner des outils aux différents champs de recherche : il s'agit alors plus de parler des sciences des données. S'il y a bien une rupture épistémologique avec l'apparition de nouveaux objets, terrains et représentations, un nouveau champ de recherche à part entière qui s'intéresse uniquement aux données, sans les processus et les contextes dont elles sont issues n'est pas un choix recommandable quand l'étude portent sur un domaine très spécifique, qui porte également le poids de ses héritages. C'est de nouveau le point de vue de Bruno Bachimont, qui s'intéresse aux transformations que peuvent créer ces transformations épistémologiques.

En s'intéressant plus particulièrement à l'épistémologie de la recherche en Droit, on constate que l'étude du contentieux n'est pas un objet mis en valeur : « *On comprend alors que la doctrine distingue plusieurs sens au mot « jurisprudence », soulignant ainsi l'ambiguïté dont il souffre de ce fait : on se réfère principalement, d'une part, à la masse indistincte des innombrables décisions, contentieux présenté souvent comme sans intérêt pour la doctrine et d'ailleurs sans portée doctrinale, et, d'autre part, à ce qui serait en quelque sorte la « vraie » jurisprudence, issue de fort peu d'arrêts mais digne d'examen, d'exégèse, de diffusion et de mémorisation. Il y a, d'une façon sous-jacente, l'idée d'une*

⁷⁸ Bachimont Bruno, « le numérique comme milieu », *Art. Cit.*

jurisprudence plus noble que d'autre »⁷⁹. Ainsi, et comme le précise Evelyne Serverin⁸⁰, les statistiques sur les juridictions et leurs productions en termes de décisions de justice, ne sont pas des données qui sont utilisées par les chercheurs en Droit. L'étude sur la masse des décisions vient transformer les manières de faire : « *Il est possible de faire le parallèle avec le temps où les seuls arrêts de la Cour de cassation connus des juristes étaient ceux sélectionnés par les universitaires pour être commentés dans les revues et où la portée normative était celle indiquée par ces commentaires. En ce temps, il était généralement considéré que la doctrine participait à la constitution de la jurisprudence: la justice prédictive jouera-t-elle un tel rôle pour les décisions du fond ?* »⁸¹

En ce sens, on comprend mieux les propos de Danièle Bourcier, qui propose de revoir l'épistémologie de la recherche en Droit : « *La science du Droit doit être distinguée de la Doctrine juridique, qui fait appel à l'opinion et aux commentaires des jurisconsultes sur le Droit positif.* »⁸². La science du Droit prend comme objet de recherche le Droit, et s'intéresse à tout ce qui le constitue : la Doctrine est un objet au même titre que les décisions de justice. La numérisation de ces dernières, et l'apport des outils statistiques permettent aujourd'hui de les exploiter. Si l'on garde également en tête le principe de construction de la jurisprudence proposé par Evelyne Serverin, on comprend plus facilement comment la recherche en Droit est associée à la Doctrine, mais c'est précisément en décrivant ce fonctionnement que l'on se trouve dans la science du Droit. L'utilisation des outils numériques vient renforcer la nécessité de dépasser la Doctrine comme fondement de la recherche en Droit, en permettant de rechercher des corrélations, des effets, qui n'étaient, autrement, pas observables : « *Lorsque la jurisprudence peut être ainsi appréhendée sous sa forme massive de contentieux, on pourrait craindre un nivellement de l'observation sous le flot des écrits et des informations ; la globalité de la perspective permet au contraire de porter un regard plus fin sur la réalité* »⁸³. Le travail de recherche en Droit se complète alors de l'étude de la production.

⁷⁹ Frisson-Roche, Marie-Anne, *Art. Cit.*

⁸⁰ « La statistique des juridictions judiciaires, pourtant la plus ancienne d'Europe (1827 pour la justice pénale, 1831 pour la justice civile et commerciale), est née et s'est développée en marge du champ documentaire exploré par les juristes, sans parvenir à s'imposer comme une donnée nécessaire de la recherche », Serverin, Evelyne, « Juridiction et jurisprudence : deux aspects des activités de justice », *Art. Cit.*

⁸¹ Deumier, Pascale, *Art. Cit.*

⁸² Bourcier, Danièle, « À propos des fondements épistémologiques d'une science du Droit », *Art. Cit.*

⁸³ Frisson-Roche, Marie-Anne, *Ibid.*

En s'inspirant des résultats que l'on a pu tirer de l'analyse des considérants de la justice administrative, nous avons pu trouver, et observer, certaines jurisprudences dans le temps et l'espace. En ce sens, il est possible d'avoir une vision macroscopique sur le fonctionnement individuelle de celles-ci : le nombre de mentions de la jurisprudence dans le temps, la proportion de son utilisation par les juridictions, le nombre d'évolutions de la jurisprudence et les date corrélées... l'ensemble de ces informations qui n'étaient pas connues auparavant pour la grande majorité des jurisprudences, sont maintenant des connaissances que l'on peut prendre en compte en étudiant une jurisprudence en particulier⁸⁴. Un savoir qui peut être également utilisé par les magistrats et permettre de re-penser le fonctionnement des bases de données à leurs dispositions. En introduisant ces nouveaux outils, nous sommes passés d'une connaissance uniquement accessible par un effort de diffusion des juristes et des juridictions supérieures, à une connaissance accessible sur une plateforme, souvent grâce à des moteurs de recherche. Nous avons maintenant accès à des statistiques d'utilisations des jurisprudences, ce qui peut amener des transformations des usages, ou du moins des transformations cognitives au sein des pratiques des magistrats : *« l'introduction de bases de données jurisprudentielles depuis les années 2000 a considérablement modifié le travail du juge. Par ailleurs, aux Etats-Unis comme dans d'autres pays, la justice pénale utilise des outils statistiques pour apporter une expertise au juge dans ce qui est appelé l'Evidence-Based Sentencing »*⁸⁵.

Le déploiement de nouvelles architectures, la production de nouvelles connaissances et de nouveaux savoirs, ainsi que la transformation cognitive qu'impliquent ces changements amènent ainsi des conséquences sur les métiers du Droit. Certains effets se font déjà ressentir⁸⁶, d'autres sont en voie de modifier les perceptions des acteurs du Droit sur leur pratiques, et de modifier les usages en cours⁸⁷. Ces conséquences ne sont pas ignorées par les magistrats du Conseil d'État ou de la Cour de Cassation, qui se questionnent sur l'influence du numérique au sein de la Justice, et en particulier sur les enjeux qui sont face à eux. La question du pouvoir

⁸⁴ « Renforçant la distinction entre contentieux, soit l'ensemble de la production des juridictions, et jurisprudence, les techniques de *data mining* pourraient en revanche rendre compte des pratiques habituelles d'un point de vue quantitatif : on saura alors ce qui est la tendance majoritaire pour un certain type d'affaires. Une telle approche statistique présente l'intérêt d'ouvrir des perspectives nouvelles qui peuvent s'avérer complémentaires d'une démarche juridique plus classique. ». Sayn, Isabelle, Clément, Marc, « L'utilisation de l'intelligence artificielle dans le système judiciaire » in Daniel Le Métayer (dir.), *nsp*, 2022

⁸⁵ Clément, Marc, *Art. Cit.*

⁸⁶ *Ibid*

⁸⁷ Frison-Roche, Marie-Anne, *Art. Cit.*

normatif de ces outils - et en particulier des modèles d'intelligences artificielles - questionnent ces instances, qui cherchent à mesurer les effets qu'ils vont amener dans leur mise en place progressive.

3. Enjeux du Droit par le prisme des outils numériques

La cour de Cassation, les pouvoirs publics, les acteurs privés ou encore le Conseil d'État se questionnent sur l'utilisation des outils numériques dans les pratiques juridiques. Les modèles d'intelligences artificielles sont particulièrement sujets à interrogations : qu'elles soient craintes ou adulées, elles polarisent les avis et sont souvent le sujet de fantasmes technico-gestionnaires. Le manque de connaissances relatif du fonctionnement de ces outils, mais surtout de leurs pré-requis, sont souvent à l'origine de ces réactions, qui empêchent souvent de considérer ces techniques comme une technologie au même titre qu'une base de données. Ce sont surtout les potentiels présumés qui sont au centre des esprits, et qui invitent à prendre en comptes les enjeux inhérents à l'utilisation croissante d'outils numériques au sein des professionnels et des chercheur.ses du Droit. En particulier plusieurs sujets sont en causes : l'automatisation des processus juridiques (de la Justice), le potentiel normatif de ces outils (sur l'influence des juridictions supérieure), les perspectives de marché, ou encore la mesure des effets de l'évolution de la loi.

Tout d'abord, l'analyse statistique des jurisprudences nécessite une publication massive des décisions de justice, et de manière qualitative, c'est-à-dire annotée. Dans notre cas, le manque d'annotation limite très fortement les potentiels d'utilisation du texte d'une décision, notamment car cela ne permet pas de contrôler finement les « outputs », c'est-à-dire les résultats des traitement numériques sur le texte. La quantité de décisions produites chaque année, ainsi que le nombre de jurisprudence et leurs spécificités rendent le travail d'annotation particulièrement laborieux. D'autant plus que cela nécessite un regard juridique pour apporter de l'information qualitative au texte « brut ». Si certaines initiatives existent ou ont existé, elles n'ont pour l'instant pas abouti⁸⁸. Il n'est donc pas encore possible de pouvoir entraîner des programmes à reconnaître automatiquement les jurisprudences à partir du cas d'espèce. À l'exception du calcul de similarité, qui implique un manque de finesse pour réellement créer un modèle de

⁸⁸ En particulier les travaux commencés par la base de données DATAJUST.

prédiction. De plus, il n'est pas prouvé que ces outils auraient un effet substantiel sur la prise de décision des juges, et à reconsidérer la production de la jurisprudence. L'effet normatif que produit la Cour de Cassation semble, encore, plus à même de contrôler l'énonciation de la règle de Droit, par le fonctionnement hiérarchique des instances judiciaires, quand bien même les magistrats restent autonomes dans la prise de décision.

Même s'il est possible de considérer que l'on peut quantifier l'usage de telle ou telle jurisprudence en fonction du cas d'espèce, on ne peut pas envisager que les juridictions supérieures n'émettent pas un avis si le procès arrive jusqu'à elle. La mise à jour des modèles avec ces avis devrait entraîner le même effet de production de la norme sur les magistrat.es. La seule possibilité de transformer le fonctionnement de la jurisprudence est donc la connaissance et l'utilisation par les magistrat.es des décisions prises par leurs pairs, sur les décisions de justice qui ne sont pas arrivées jusqu'à la Cour de Cassation, ou le Conseil d'État. Fonctionnement qui existe déjà en pratique, mais dont l'innovation technique permettrait un travail plus rapide pour les professionnels du Droit, sans pour autant rendre le travail efficace. L'exemple des barèmes proposé par Isabelle Sayn est un parallèle intéressant avec les modèles de prédictions : « *Les nombreux entretiens réalisés avec des magistrats du premier degré à l'occasion d'une recherche sur l'usage des barèmes dans les juridictions montrent que ces démarches ne sont jamais conçues comme un succédané de jurisprudence, d'abord parce qu'elles laissent chaque magistrat libre d'y adhérer ou pas, ensuite parce qu'elles ne prétendent en aucun cas produire une interprétation des règles de Droit, sauf peut-être à combler à l'occasion des lacunes repérées dans la jurisprudence de la Cour de cassation s'agissant de point de Droit qui ne lui ont pas été soumis et qui se répètent devant la juridiction* »⁸⁹.

En parallèle des professionnel.les du Droit, l'utilisation des outils numériques amènent un ensemble d'acteurs privés qui souhaitent s'insérer sur ce que Boris Barraud appelle les « marchés du Droit » : « *L'intelligence artificielle, sitôt qu'elle côtoie le Droit, engendre l'apparition de nouveaux marchés à mesure que des demandes auxquelles il était auparavant impossible de répondre peuvent désormais être satisfaites grâce à la technologie. [...] Ce sont des legal start-up qui, saisissant l'opportunité de nouveaux marchés juridiques encore peu explorés, proposent des technologies subrogeant des services automatiques à des services humains.*»⁹⁰. Ces legal-tech peuvent exploiter de nouveaux objets à même de produire des outils et des connaissances qui peuvent servir

⁸⁹ Sayn, Isabelle. Justice et numérique. Quelles (r)évolutions? p. 72.

⁹⁰ Barraud, Boris, *Art. Cit.*

de marchandises auprès des institutions juridiques, et en particulier des cabinets d'avocats. La création d'un tel marché repose également sur la croyance selon laquelle les technologies numériques peuvent permettre d'apporter des informations sur les issues d'un procès, c'est-à-dire de prédire les décisions des juges en fonction des éléments du procès. Nous avons vu que cette théorie en plus d'être hasardeuse nécessite des travaux conséquents au préalable. Ces entreprises s'orientent donc sur un marché spécifique, supposé plus simple et plus rémunérateur : les décisions de justice chiffrées, qui portent sur des contentieux indemnitaires.

Plus simple dans la mesure où elles nécessitent moins d'expertises pour l'annotation, qui peut être éventuellement automatisable. C'est ce qu'essaye de prouver un autre projet d'e-juris qui cherche à repérer et classer automatiquement les informations chiffrées des décisions de justice portant sur les divorces. Les résultats, bien qu'intéressants, montrent particulièrement l'importance d'une quantité élevée de décisions annotées pour fonctionner, sans pour autant garantir un bon fonctionnement en fonction des informations que l'on cherche à prédire : il est plus simple de retrouver les dates de mariage et de divorce que les montants demandés par les parties.

Il existe ainsi un enjeu non négligeable pour les acteurs privés d'investir un nouveau marché qui est permis par la numérisation des documents et la mise en place d'outils informatiques pour les traiter. C'est également un enjeu pour l'Etat et les cours supérieures, pour essayer de quantifier, ou de prédire, l'impact des réformes de la loi sur les groupes sociaux. Pouvoir observer les usages de la loi par les statistiques, c'est donner un outil qui permet de voir les utilisations qui en sont faites, c'est-à-dire les jurisprudences qui sont construites dans la mise à l'épreuve de la loi avec les faits : « *le législateur lance une loi selon l'idée qu'il a de sa réception, mais il ne peut prendre la mesure de tous ses effets, notamment de ses effets pervers. Il est donc sain de connaître le plus rapidement possible l'usage qu'en fait la jurisprudence, par le premier choc nécessairement déformant que subit la norme dans sa rencontre avec le fait et le juge. Cela permettra au législateur au besoin de combattre l'acclimatation de la norme par la jurisprudence massive, s'il la juge illégitime* »⁹¹. C'est également l'approche quantitative qui peut permettre d'observer comment des contentieux particuliers peuvent se créer, en dehors de toute règle de Droit pour l'encadrer, et de pouvoir quantifier ces réalités juridiques.

⁹¹ Frison-Roche, Marie-Anne, *Art. Cit.*

L'utilisation de technologie numérique modifie le rapport des scientifiques au Droit, le regard des professionnel.les sur leur profession, ainsi que les potentiels pour une diversité d'acteurs et d'actrice de la société. En utilisant les nouvelles ressources que sont les données issues des décisions de justice, de nouveaux terrains sont à même d'être traités, et ainsi des questions qui n'auraient pas pu être posées autrement, sont à même d'être débattues. Ce mouvement transforme lentement le travail de recherche ainsi que le regard qui est porté sur le Droit et le fonctionnement de la justice par les professionnel.les du Droit, mais aussi par les acteurs qui se trouvent en appuie au fonctionnement du Droit. Force est de constater que ces évolutions questionnent également les institutions sur la manière dont les règles de Droit sont produites, et sur la question du remplacement des magistrats par des ordinateurs omniscients. C'est oublier la manière dont les jurisprudences se construisent aujourd'hui, et dépasser le travail des magistrats eux-mêmes, qui font vivre les jurisprudences, tout autant qu'ils les créent, si l'on considère qu'ils et elles sont toujours indépendant.es en théories des cours supérieures. L'application d'une norme dépend alors du raisonnement juridique que se fait le ou la magistrat.e sur un litige, qui évolue également dans le temps et dans les litiges. Ce pouvoir que représente l'indépendance des magistrat.es, qui remet ainsi la jurisprudence constamment à l'épreuve des faits, ainsi que la difficulté de tirer des données qualitatives des décisions de justice sans annotation, rend aujourd'hui le travail d'automatisation de la prise de décision particulièrement hasardeux.

Conclusion

In fine, qu'avons-nous pu montrer au sein de ce mémoire ? La diffusion des décisions de justice qui résulte d'un processus historique de publication des documents administratifs, permet de créer de nouvelles approches au sein de la recherche en Droit. Le travail de co-construction des corpus est nécessaire à un travail de qualité données, même s'il n'est pas suffisant : l'expertise juridique permet aux chercheurs de savoir quelles données sont à même d'être extraites des documents, et les formes de traitements auxquelles elles peuvent être soumises. Ce travail d'extraction au sein du texte requiert de construire des frontières textuelles, afin de pouvoir constituer un corpus de document que l'on souhaite exploiter. Dans notre cas, nous avons opéré un ensemble de sélection, de transformation et de tri afin de pouvoir au mieux repérer et mettre en comparaison les jurisprudences et les considérants qui les citent. Sans l'apport d'un côté de l'expertise juridique, et de l'autre de l'expertise informatique, ce travail n'aurait pas le même sens, et manquerait d'une part de contextes juridiques dans l'extraction, de l'autre d'une expertise technique pour mettre en place cette extraction. C'est dans cette hybridation que l'on peut, à mon sens, reconstruire le plus de sens dans l'analyse, car les résultats dépendent très largement des conditions d'extractions et de traitements des données : « *On comprend alors pourquoi les études sur les big data en font trop ou pas assez : elles vont trop loin car elles perdent l'humanité des faits qu'elles étudient. Si bien que les résultats brandis sont bien des propriétés des modèles d'analyse, mais ne nous apprennent que fort peu sur la réalité dont les données sont issues. Et donc elles ne vont pas assez loin : pour tenir le programme qu'elles se sont donné, il leur faudrait aller au bout de l'interprétation et comprendre en quoi les analyses menées permettent de retrouver l'humanité des faits étudiés et ainsi avancer dans la connaissance et la compréhension de la culture.* »⁹² C'est précisément du fait que les résultats sont en parties produits par les choix et les données, qu'il réside encore des limites à un triomphe de la science du Droit : le manque d'annotations, les nombreux cas d'exception dans les données, ou encore le manque de « raffinage »⁹³ des données que l'on a extraites sont autant de limites aux différents travaux qui peuvent être réalisés à ce stade sur les décisions de justice.

⁹² Bachimont, Bruno, « Le nominalisme et la culture », *Art. Cit.*

⁹³ Qui nécessite à mon avis de revenir assez régulièrement sur les processus d'extraction : après que l'on ait pu observer certaines limites, éliminer les individus sur- et sous-représentés,...

Cependant, cette limitation ne doit pas nous éloigner des résultats qui ont pu être tirés de notre travail : il s'agit en effet d'un domaine de connaissance nouveau pour l'étude du Droit, qui se caractérise par la production de connaissances à partir du contentieux, et la possibilité de se pencher sur de nouveaux terrains. Les données sont ainsi un enjeu majeur à la fois pour la recherche en Droit, mais aussi pour les praticiens et les acteurs des marchés du Droit : l'accès aux données permet de donner un regard sur les pratiques et les usages, tout en étant lié au réseau d'interactions institutionnelles, qui contribuent à la production des décisions de justice, qui pèsent sur les résultats et les analyses. Il ne s'agit pas de dire que les pratiques ne sont pas vouées à évoluer, bien au contraire, mais de comprendre que la production de nouvelles connaissances et de nouveaux outils n'amène pas nécessairement une révolution dans la manière de faire le Droit, de produire des décisions de justice, de « gouverner » l'appareil juridique. Il s'agit plus en effet d'une lente évolution des questions de recherche, des outils utilisés par les professionnel.les, et des acteurs qui interviennent sur le marché du Droit, transformé en partie par la production des données.

* * *

« La justice, ce n'est pas seulement appliquer la loi. C'est réussir à ce que, par une opération très ritualisée et très complexe, on arrive à se mettre d'accord sur l'issue d'un certain litige. Et tout ça prend du temps. C'est pour cela qu'un procès, c'est long et compliqué, et avec toutes sortes de ritualisations collectives. Si l'on oublie toute la dimension anthropologique du Droit pour ne conserver que sa dimension algorithmique, ce qui est un rêve formaliste pour certains juristes, on rate le but qui est de rendre justice »⁹⁴. Cette citation de Jean Lassègue et d'Antoine Garapon nous permet de remettre d'abord en contexte ce que l'on a pu étudier dans notre mémoire : il s'agit avant tout de comprendre le processus de transformation de l'information qui s'opère avec les traitements numériques que subissent les documents, afin de pouvoir au mieux reconstruire du sens avec les résultats. On y comprend également tout le poids de l'institution, des rituels, du vocabulaire, qui sont autant d'éléments à prendre en compte dans la mise en calcul du Droit. Statuer sur une décision, c'est tout autant écrire le Droit que rendre justice, ce qui implique d'une part un processus de réflexion, de compréhension des enjeux et d'application des règles de Droit adaptées, et d'autre part

⁹⁴ Garapon, Antoine, et Jean Lassègue. Justice digitale: Révolution graphique et rupture anthropologique. Presses Universitaires de France, 2018.

de réaliser le procès, qui fait exister une décision fondée sur un ensemble complexe de règles à une décision qui porte et représente le poids de l'Etat.

Cette citation implique aussi que considérer le Droit et les décisions de justice uniquement en tant qu'inputs d'un programme, c'est oublier l'ensemble des règles et rites qui concourent à la production de ces décisions. Pour continuer dans cette perspective, on peut s'intéresser également à la cette citation de Pascale Deumier : « *Les enjeux de la justice prédictive sont désormais bien connus . D'un côté, elle fait miroiter le développement d'un nouveau marché numérique, ce qui a motivé les amendements ayant introduit l'open data des décisions ; la transparence de la justice à l'égard du citoyen ; l'harmonisation de l'application du Droit sur le territoire ; le désencombrement des juridictions par l'augmentation des résolutions amiables ou le renoncement à une action vouée à l'échec. De l'autre, elle fait redouter la pression exercée sur la prise de décision des juges ; l'oubli de la spécificité des cas ; la possibilité de manipuler les algorithmes et les statistiques ; la transformation d'une part de la justice en marché* »⁹⁵. L'apport des humanités numériques est d'ainsi prendre en compte cette « humanité » de la données, afin de la re-construire, ou de la prendre en compte en amont et en aval de la production des résultats. C'est d'ainsi comprendre que l'extraction, la production et la manipulation de données sont en soi un fait social durkheimien dans le sens où elles transforment le rapport de l'Homme à la société - et en particulier à la Science - et d'en comprendre les milieux spécifiques qu'elles transforment.

⁹⁵ Deumier, Pascale, *Art. Cit.*

Bibliographie

Bachimont, Bruno. Le nominalisme et la culture : questions posées par les enjeux du numérique. p. 8.

---. « Le numérique comme milieu : enjeux épistémologiques et phénoménologiques. : Principes pour une science des données ». *Interfaces numériques*, vol. 4, n° 3, 3, décembre 2017, p. 402-402. www.unilim.fr, <https://doi.org/10.25965/interfaces-numeriques.386>.

Barraud, Boris. « Le Droit en datas : comment l'intelligence artificielle redessine le monde juridique (première partie : La dictature des algorithmes ou l'intelligence artificielle à la source du Droit) ». *Revue Lamy Droit de l'immatériel*, 2019. HAL Archives Ouvertes, <https://hal.archives-ouvertes.fr/hal-02073660>.

Barthe, Emmanuel. « L'intelligence artificielle et le Droit ». *I2D - Information, donnees documents*, vol. 54, n° 2, juillet 2017, p. 23-24.

Barua, Josh. « Word Embeddings Versus Bag-of-Words: The Curious Case of Recommender Systems ». *The Startup*, 25 août 2020, <https://medium.com/swlh/word-embeddings-versus-bag-of-words-the-curious-case-of-recommender-systems-6ac1604d4424>.

Béroujon, Christiane, Pour une analyse empirique des relations entre contentieux et jurisprudence, RTD Civ., 1993 p.94

Bories, Serge. « L'informatisation des données judiciaires et doctrinales : une contribution à la connaissance et à la recherche juridiques ». *Documentaliste-Sciences de l'Information*, vol. 40, n° 4, 2003, p. 272-79.

Bourcier, Danièle. *À propos des fondements épistémologiques d'une science du Droit*. Presses Universitaires de France, 2007. www-cairn-info.bibelec.univ-lyon2.fr, <http://www.cairn.info/quelles-perspectives-pour-la-recherche-juridique--9782130558989-page-69.htm>.

Bourcier, Danièle, et Primavera De Filippi. « Transparency of algorithms in a context of open data: which status for learning data? » *Revue française d'administration publique*, vol. 167, n° 3, 2018, p. 525-37.

Bouzidi, Laïd, et Sabrina Boulesnane. « Les humanités numériques ». *Les Cahiers du numérique*, vol. 13, n° 3, 2017, p. 19-38.

Bruxelles, Sylvie, et Evelyne Serverin. « Du judiciaire au juridique : un procès d'avortement dans les revues de jurisprudence ». *Langages*, vol. 12, n° 53, 1979, p. 51-65. www.persee.fr, <https://doi.org/10.3406/lgge.1979.1812>.

Carbou, Guillaume. « Analyser les textes à l'ère des humanités numériques ». *Les Cahiers du numérique*, vol. 13, n° 3, 2017, p. 91-114.

Cardon, Dominique. *A quoi rêvent les algorithmes : nos vies à l'heure des big data*. Seuil, 2015.

Clément, Marc, « Algorithmes au service du juge administratif : peut-on en rester maître ? », *AJDA*, 2017 p.2453

Cottin, Marianne, « Jurisprudence et contentieux, une (r)évolution à attendre ? », *Justice et numérique. Quelles (r)évolutions ?*, Lexis-Nexis, p.22-25

« Cour de discipline budgétaire et financière ». *Wikipédia*, 2 octobre 2021. *Wikipedia*, [https://fr.wikipedia.org/w/index.php?title=Cour de discipline budg%C3%A9taire et financi%C3%A8re&oldid=186811489](https://fr.wikipedia.org/w/index.php?title=Cour_de_discipline_budg%C3%A9taire_et_financi%C3%A8re&oldid=186811489).

« DataJust : violer la loi sous couvert d'expérimentation ». *La Quadrature du Net*, 23 décembre 2021, <https://www.laquadrature.net/2021/12/23/datajust-violer-la-loi-sous-couvert-dexperimentation/>.

Deumier, Pascale. « La justice prédictive et les sources du Droit : la jurisprudence du fond ». *Archives de philosophie du Droit*, vol. 60, n° 1, 2018, p. 49-66.

Fathisalout-Bollon, Motahareh, et al. « Décret DataJust : un algorithme pour indemniser les dommages corporels ? » *Fenêtre sur Cour*, n° 5, octobre 2020, p. 26.

FRISON-ROCHE, Marie-Anne, *La jurisprudence massive*, en collaboration avec Serge BORIES, D.1993, chron., p.287 s.

Garapon, Antoine, et Jean Lassègue. *Justice digitale: Révolution graphique et rupture anthropologique*. Presses Universitaires de France, 2018.

Goulet, Jean. « Revalorisation du Droit et jurimétrie ». *Les Cahiers de Droit*, vol. 9, n° 1, 1967, p. 9-36. www.erudit.org, <https://doi.org/10.7202/1004340ar>.

Hourdeaux, Jérôme. « La justice se prépare à l'arrivée des algorithmes ». *Mediapart*, <http://www.mediapart.fr/journal/france/020119/la-justice-se-prepare-l-arrivee-des-algorithmes>. Consulté le 17 mai 2022.

Jeammaud, Antoine, Serverin, Evelyne, « *Concevoir l'espace jurisprudentiel* », *RTD civ.*, 1993, p. 91

Latour, Bruno, *La fabrique du Droit*, La Découverte, 2004, www-cairn-info.bibelec.univ-lyon2.fr, <https://doi.org/10.3917/dec.latou.2004.03>.

Lemercier, Claire. « Que faire du matériau juridique en sciences sociales ? » *Droit et sciences sociales*, <https://Droitscisoc.hypotheses.org/509>. Consulté le 11 mai 2022.

« Loi pour une République numérique ». *Wikipédia*, 3 mai 2022. *Wikipedia*, [https://fr.wikipedia.org/w/index.php?title=Loi pour une R%C3%A9publique num%C3%A9rique&oldid=193367622](https://fr.wikipedia.org/w/index.php?title=Loi_pour_une_R%C3%A9publique_num%C3%A9rique&oldid=193367622).

Longhi, Julien. « Humanités, numérique : des corpus au sens, du sens aux corpus ». *Questions de communication*, vol. 31, n° 1, septembre 2017, p. 7-17.

Poibeau, Thierry. « Le traitement automatique des langues pour les sciences sociales: Quelques éléments de réflexion à partir d'expériences récentes ». *Réseaux*, vol. n° 188, n° 6, 2014, p. 25. *DOI.org (Crossref)*, <https://doi.org/10.3917/res.188.0025>.

« Recueil Lebon ». *Wikipédia*, 25 février 2022. *Wikipedia*, [https://fr.wikipedia.org/w/index.php?title=Recueil Lebon&oldid=191408931](https://fr.wikipedia.org/w/index.php?title=Recueil_Lebon&oldid=191408931).

Rivollier, Vincent, et al. « Le retrait de DataJust, ou la fausse défaite des barèmes ». *Recueil Dalloz*, n° 09, mars 2022, p. 467.

Sayn, Isabelle. Justice et numérique. Quelles (r)évolutions? p. 72.

Sayn, Isabelle, Clément, Marc, « L'utilisation de l'intelligence artificielle dans le système judiciaire » in Daniel Le Métayer (dir.), *nsp*, 2022

Slama, Rémi. « La Chine annonce avoir développé une intelligence artificielle capable de remplacer les juges ». *Laboratoire de cyberjustice*, <https://www.cyberjustice.ca/2022/02/01/la-chine-annonce-avoir-developpe-une-intelligence-artificielle-capable-de-remplacer-les-juges/>. Consulté le 29 juin 2022.

Serverin, Évelyne. « De l'informatique juridique aux services de justice prédictive, la longue route de l'accès du public aux décisions de justice dématérialisées ». *Archives de philosophie du Droit*, vol. 60, n° 1, 2018, p. 23-47.

---. « Juridiction et jurisprudence : deux aspects des activités de justice ». *Droit et Société*, vol. 25, n° 1, 1993, p. 339-49. *www.persee.fr*, <https://doi.org/10.3406/dreso.1993.1229>.

---. *Sociologie du Droit / Évelyne Serverin*. 2000. *gallica.bnf.fr*, <https://gallica.bnf.fr/ark:/12148/bpt6k4813537p>.

Le code source des programmes produits dans le cadre du stage et du mémoire sont disponibles à cette adresse : <https://gitlab.huma-num.fr/amassein/e-Juris>

Table des Matières

L'état du Droit	1
Expérience de reconnaissance automatique des jurisprudences de la justice administrative	1
Remerciements	2
Introduction	3
I. De la décision de justice aux données juridiques : constituer le corpus	12
A. Historique et questionnements autour de la numérisation des documents juridiques	13
B. Les décisions de justice : de la décision aux données.	16
1. <i>La base de données JADE et les décisions de justice administrative</i>	16
2. <i>Les métadonnées</i>	19
3. <i>Le texte de la décision</i>	21
C. Comment utiliser ces données ? Sélectionner l'information	24
1. <i>Sur les métadonnées</i>	24
2. <i>Sur le texte des décisions</i>	27
II. Des données juridiques à la science du Droit : l'apport du numérique et le poids des choix	30
A. Traiter le corpus pour le rendre analysable : vectoriser le texte	31
1. <i>Rendre le texte comparable</i>	31
2. <i>Automatiser la recherche des jurisprudences</i>	33
B. Exemples d'analyses	36
1. <i>Étude sur la concentration des considérants sur quelques jurisprudences.</i>	37
2. <i>Étude d'un cas particulier de jurisprudence : limites du travail d'analyse</i>	43
C. Une révolution (épistémologique) du Droit ?	48
1. <i>De « nouvelles » connaissances ?</i>	49
2. <i>Faire évoluer l'étude des jurisprudences et des contentieux</i>	52
3. <i>Enjeux du Droit par le prisme des outils numériques</i>	55
Conclusion	59
Bibliographie	62
Table des Matières	65
Annexes	67

1. Graphiques et documents	68
1. Répartitions par années des décisions de justice administrative	68
2. Exemple d'une décision de justice au format XML	69
3. Exemple d'une décision de justice au format papier	70
4. Graphique des types de recours qui sont mentionnés plus de 50 fois au sein des décisions de justice	71
5. Graphique des types de recours par année sur les recours mentionnés plus de 50 fois au sein des décisions de justice	71
6. Graphique des solutions	72
7. Graphique des 50 jurisprudences les plus utilisées	72
8. Graphiques des références légales de la jurisprudence 35553	73
9. Graphiques des références légales de la jurisprudence 32771	74
10. Distribution des contentieux pondéré sur le nombre de considérant par jurisprudence.	75
11. Distribution des considérants associés à la jurisprudence 49511, et jurisprudence proches, en fonction de la juridiction	76
12. Distribution de l'ensemble des jurisprudences proche de 49511, en fonction de l'année et de la juridiction	77
13. Distribution des considérants par année et par juridiction pour la jurisprudence 72913 en pourcentage par année	78
2. Exemples de jurisprudences	79
1. Exemples des 7 jurisprudences les plus mentionnées	79
2. Jurisprudences 49511 et « proches »	81

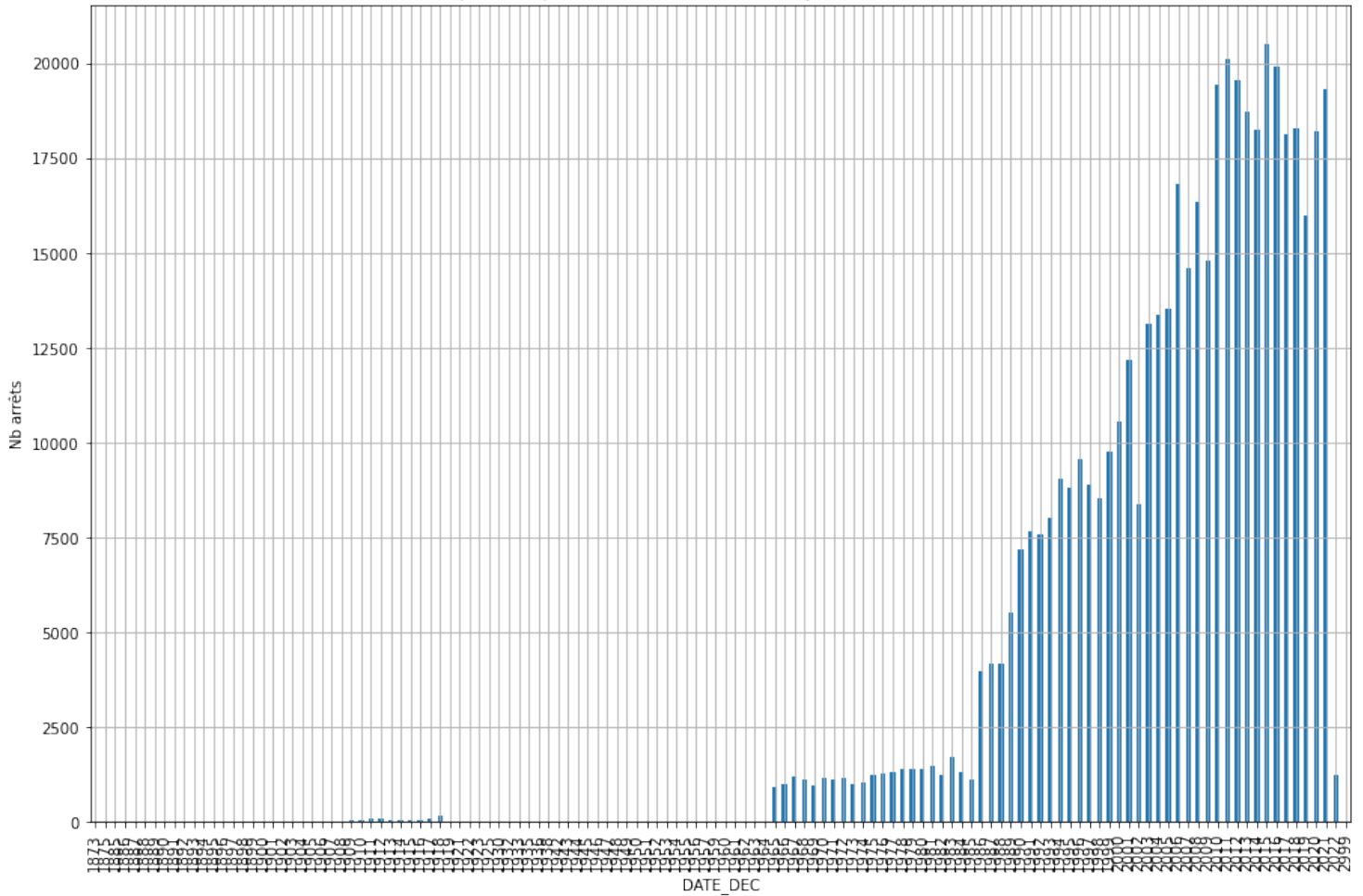
Annexes

1. Graphiques et documents	68
1. Répartitions par années des décisions de justice administrative	68
2. Exemple d'une décision de justice au format XML	69
3. Exemple d'une décision de justice au format papier	70
4. Graphique des types de recours qui sont mentionnés plus de 50 fois au sein des décisions de justice	71
5. Graphique des types de recours par année sur les recours mentionnés plus de 50 fois au sein des décisions de justice	71
6. Graphique des solutions	72
7. Graphique des 50 jurisprudences les plus utilisées	72
8. Graphiques des références légales de la jurisprudence 35553	73
9. Graphiques des références légales de la jurisprudence 32771	74
10. Distribution des contentieux pondéré sur le nombre de considérant par jurisprudence.	75
11. Distribution des considérants associés à la jurisprudence 49511, et jurisprudence proches, en fonction de la juridiction	76
12. Distribution de l'ensemble des jurisprudences proche de 49511, en fonction de l'année et de la juridiction	77
13. Distribution des considérants par année et par juridiction pour la jurisprudence 72913 en pourcentage par année	78
2. Exemples de jurisprudences	79
1. Exemples des 7 jurisprudences les plus mentionnées	79
2. Jurisprudences 49511 et « proches »	81

1. Graphiques et documents

1. Répartitions par années des décisions de justice administrative

Répartition par années des décisions de justice administrative



2. Exemple d'une décision de justice au format XML

file:/Users/Aljoscha/Desktop/M2 - Stage/e-Juris_Gitlab/jade_lebon/CETATEXT000007612831.xml

```
1 <?xml version="1.0" encoding="UTF-8"?>
2 <TEXTE_JURI_ADMIN>
3   <META>
4     <META_COMMUN>
5       <ID>CETATEXT000007612831</ID>
6       <ANCIEN_ID>JFXLX1975X03X0000081158</ANCIEN_ID>
7       <ORIGINE>CETAT</ORIGINE>
8
9   <URL>texte/juri/admin/CETA/TEXT/00/00/07/61/28/CETATEXT000007612831
10  .xml</URL>
11   <NATURE>Texte</NATURE>
12   <META_COMMUN>
13   <META_SPEC>
14     <META_JURI>
15       <TITRE>Conseil d'Etat, 9 / 8 SSR, du 12 mars 1975,
16 81158, mentionné aux tables du
17 recueil Lebon </TITRE>
18       <DATE_DEC>1975-03-12</DATE_DEC>
19       <JURIDICTION>Conseil d'Etat </JURIDICTION>
20       <NUMERO>81158</NUMERO>
21       <SOLUTION>REJET Droits maintenus </SOLUTION>
22     </META_JURI>
23     <META_JURI_ADMIN>
24       <FORMATION>9 / 8 SSR</FORMATION>
25       <TYPE_REC>Plein contentieux fiscal </TYPE_REC>
26       <PUBLI_RECUEIL>B</PUBLI_RECUEIL>
27       <DEMANDEUR/>
28       <DEFENDEUR/>
29       <PRESIDENT>M. Rain</PRESIDENT>
30       <AVOCATS/>
31       <RAPPORTEUR>M. Rivière</RAPPORTEUR>
32       <COMMISSAIRE_GVT>M. Schmeltz </COMMISSAIRE_GVT>
33     </META_JURI_ADMIN>
34   </META_SPEC>
35 </META>
36 <TEXTE>
37   <BLOC_TEXTUEL>
38     <CONTENU> VU LA DECISION EN DATE DU 22 NOVEMBRE 1972
39 PAR LAQUELLE LE CONSEIL D'ETAT
40 STATUANT AU CONTENTIEUX AVANT DIRE DROIT SUR LA
41 REQUETE DU SIEUR Y... X...,
42 DEMEURANT RUE ROGER Z... A LONGUENESSE
43 PAS-DE-CALAIS , ENREGISTREE SOUS LE N° 81 159
44 ET TENDANT A L'ANNULATION D'UN JUGEMENT EN DATE DU
45 5 JUIN 1970 PAR LEQUEL LE
```

1

file:/Users/Aljoscha/Desktop/M2 - Stage/e-Juris_Gitlab/jade_lebon/CETATEXT000007612831.xml

```
86 TRIBUNAL ADMINISTRATIF DE LILLE A REJETE SES
87 DEMANDES EN REDUCTION ET EN RESTITUTION
88 DES IMPOSITIONS LITIGIEUSES ; <br/> DECIDE : ARTICLE
89 1ER.- LA REQUETE SUSVISEE DU
90 SIEUR X... EST REJETEE. ARTICLE 2.- EXPEDITION DE
91 LA PRESENTE DECISION SERA
92 TRANSMISE AU MINISTRE DE L'ECONOMIE ET DES
93 FINANCES.<br/>
94 </CONTENU>
95 </BLOC_TEXTUEL>
96 <SOMMAIRE>
97   <SCT ID="8AA" TYPE="PRINCIPAL">19-06-01-01,RJ1,RJ2
98 CONTRIBUTIONS ET TAXES - TAXES SUR LE
99 CHIFFRE D'AFFAIRES ET ASSIMILES - QUESTIONS
100 COMMUNES - CHAMP D'APPLICATION DES
101 T.C.A. - Notion de producteur par tiers. </SCT>
102   <SCT ID="8BA" TYPE="PRINCIPAL">19-06-01-02,RJ1,RJ2
103 CONTRIBUTIONS ET TAXES - TAXES SUR LE
104 CHIFFRE D'AFFAIRES ET ASSIMILES - QUESTIONS
105 COMMUNES - EXEMPTIONS ET EXONERATIONS -
106 Travaux de façon exécutés pour le compte
107 d'assujettis à la T.V.A - Notion de
108 producteur par tiers. </SCT>
109   <SCT ID="8CA" TYPE="PRINCIPAL">19-06-02-01,RJ1,RJ2
110 CONTRIBUTIONS ET TAXES - TAXES SUR LE
111 CHIFFRE D'AFFAIRES ET ASSIMILES - T.V.A. -
112 PERSONNES ET AFFAIRES TAXABLES -Notion de
113 producteur par tiers. </SCT>
114   <ANA ID="9AA">19-06-01-01, 19-06-01-02, 19-06-02-01
115 Entreprise installant des moteurs
116 sur des péniches. Les marinières n'imposaient à
117 l'entreprise aucune directive autre
118 que celles qui sont habituelles de client à
119 fournisseur. L'entreprise vendait
120 elle-même aux marinières les moteurs à installer et
121 leur fournissait les autres
122 matériels nécessaires à l'exécution des travaux ;
123 lorsque les marinières payaient
124 directement les moteurs aux fabricants, l
125 'entreprise leur servait d'intermédiaire.
126 Dans ces conditions les marinières n'ont pas agi
127 comme producteurs par tiers et
128 l'entreprise était assujettie aux taxes sur le
129 chiffre d'affaires pour ses
130 opérations d'installation [1] [2]. </ANA>
131 </SOMMAIRE>
```

4

3. Exemple d'une décision de justice au format papier

Firefox

<https://www.conseil-etat.fr/arianeweb/>

Conseil d'État

N° 81158

ECLI:FR:CESSR:1975:81158.19750312

Mentionné au tables du recueil Lebon

Section du Contentieux

M. Rain, président
M. Rivière, rapporteur
M. Schmeltz, commissaire du gouvernement

Lecture du 12 mars 1975

REPUBLIQUE FRANCAISE
AU NOM DU PEUPLE FRANCAIS

VU LA DECISION EN DATE DU 22 NOVEMBRE 1972 PAR LAQUELLE LE CONSEIL D'ETAT STATUANT AU CONTENTIEUX AVANT DIRE DROIT SUR LA REQUETE DU SIEUR Y... X..., DEMEURANT RUE ROGER Z... A LONGUENESSE PAS-DE-CALAIS , ENREGISTREE SOUS LE N° 81 159 ET TENDANT A L'ANNULATION D'UN JUGEMENT EN DATE DU 5 JUIN 1970 PAR LEQUEL LE TRIBUNAL ADMINISTRATIF DE LILLE A REJETE : 1° SA DEMANDE EN REDUCTION DES DROITS ET PENALITES AUXQUELS IL A ETE ASSUJETTI PAR UN AVIS DE MISE EN RECouvreMENT DECERNE A SON ENCONTRE LE 22 MAI 1964 AU TITRE DE LA TAXE SUR LES PRESTATIONS DE SERVICES POUR LA PERIODE DU 1ER DECEMBRE 1959 AU 31 DECEMBRE 1960 ; 2° SA DEMANDE EN RESTITUTION DES DROITS PAYES SPONTANEMENT PAR LUI AU TITRE DE LA TAXE SUR LES PRESTATIONS DE SERVICES ET DE LA TAXE LOCALE SUR LE CHIFFRE D'AFFAIRES POUR LA PERIODE DU 1ER MARS 1959 AU 31 DECEMBRE 1960 ;

A ORDONNE UN SUPPLEMENT D'INSTRUCTION AUX FINS D'ETABLIR SI LES TRAVAUX DE FACON EXECUTES PAR LE REQUERANT PENDANT LA PERIODE ALLANT DU 26 OCTOBRE 1959 AU 31 DECEMBRE 1960 ONT ETE EXECUTES POUR LE COMPTE DE MAITRES D'OEUVRE ASSUJETTIS A LA TAXE SUR LA VALEUR AJOUTEE ; VU LE CODE GENERAL DES IMPOTS ; VU L'ORDONNANCE DU 31 JUILLET 1945 ET LE DECRET DU 30 SEPTEMBRE 1953 ;

CONSIDERANT QUE, PAR DECISION EN DATE DU 22 NOVEMBRE 1972, LE CONSEIL D'ETAT A ORDONNE UN SUPPLEMENT D'INSTRUCTION AUX FINS DE PERMETTRE AU REQUERANT D'ETABLIR QUE LES TRAVAUX EXECUTES PAR LUI PENDANT LA PERIODE A LAQUELLE SE RAPPORTENT LES IMPOSITIONS CONTESTEES ONT ETE EXECUTES, COMME IL LE SOUTIENT, POUR LE COMPTE DE MAITRES D'OEUVRE ASSUJETTIS A LA TAXE SUR LA VALEUR AJOUTEE ;

CONSIDERANT QU'EN VERTU DES ARTICLES 271-20° ET 1575-2-27° DU CODE GENERAL DES IMPOTS, DANS LEUR REDACTION APPLICABLE AU COURS DE LADITE PERIODE, SONT EXEMPTES DE LA TAXE SUR LES PRESTATIONS DE SERVICES ET DE LA TAXE LOCALE SUR LE CHIFFRE D'AFFAIRE LES TRAVAUX OU OPERATIONS DE FACON "EXECUTES POUR LE COMPTE D'ASSUJETTIS A LA TAXE SUR LA VALEUR AJOUTEE" ; QU'EN APPLICATION DE L'ARTICLE 263 DU MEME CODE, LES PRODUCTEURS SONT ASSUJETTIS A LA TAXE SUR LA VALEUR AJOUTEE, ET QU'AUX TERMES DE L'ARTICLE 264, "PAR PRODUCTEUR, IL FAUT ENTENDRE A LES PERSONNES OU SOCIETES QUI, A TITRE PRINCIPAL OU ACCESSOIRE, FABRIQUENT DES PRODUITS OU LEUR FONT SUBIR DES FACONS COMPORTANT OU NON L'EMPLOI D'AUTRES MATIERES, SOIT POUR LA FABRICATION DES PRODUITS, SOIT POUR LEUR PRESENTATION COMMERCIALE :... C LES PERSONNES OU SOCIETES QUI FONT EFFECTUER PAR DES TIERS LES OPERATIONS VISEES AUX ALINEAS A ET B CI-DESSUS" ;

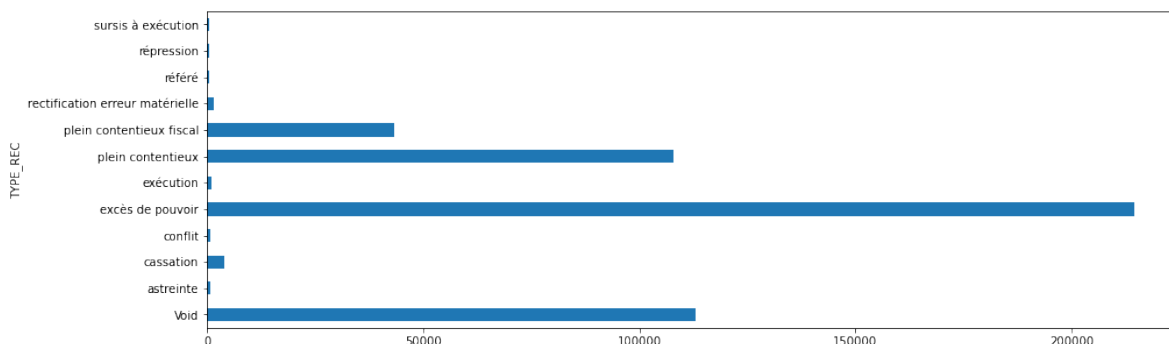
CONSIDERANT QUE LES IMPOSITIONS LITIGIEUSES ONT ETE ETABLIES A RAISON DES TRAVAUX D'INSTALLATION DE MOTEURS EFFECTUES SUR DES PENICHES PAR LE SIEUR X... QU'IL RESULTE DU SUPPLEMENT D'INSTRUCTION AUQUEL IL A ETE PROCEDE EN EXECUTION DE LA DECISION SUSVISEE QUE LES MARINIERS PROPRIETAIRES DE CES PENICHES N'ONT IMPOSE A CETTE OCCASION AU REQUERANT AUCUNE DIRECTIVE AUTRE QUE CELLES QUI SONT HABITUELLES DE CLIENT A FOURNISSEUR ; QUE LE REQUERANT RENDAIT LUI-MEME AUX MARINIERS, DANS CERTAINS CAS, LES MOTEURS A INSTALLER ET LEUR FOURNISSAIT LES AUTRES MATERIELS NECESSAIRES A L'EXECUTION DES TRAVAUX ; QUE LE SIEUR X... SERVAIT EGALEMENT D'INTERMEDIAIRE ENTRE LES MARINIERS ET LES FABRICANTS DE MOTEURS, DANS LES CAS OU LES MARINIERS ASSURAIENT DIRECTEMENT A CES DERNIERS LE PAIEMENT DE CES MOTEURS ; QU'IL SUIT DE LA QUE LES MARINIERS PROPRIETAIRES DES PENICHES NE PEUVENT ETRE REPUTES AVOIR AGI, EN L'ESPECE, COMME PRODUCTEURS PAR TIERS AU SENS DE L'ARTICLE 264 C DU CODE GENERAL DES IMPOTS ET N'ETAIENT DES LORS PAS, A CE TITRE, ASSUJETTIS A LA TAXE SUR LA VALEUR AJOUTEE ;

CONSIDERANT QU'IL RESULTE DE CE QUI PRECEDE QUE LE SIEUR X... NE POUVAIT PAS PRETENDRE, A RAISON DES OPERATIONS D'INSTALLATION DE MOTEURS AUXQUELLES IL A PROCEDE, BENEFICIER DE

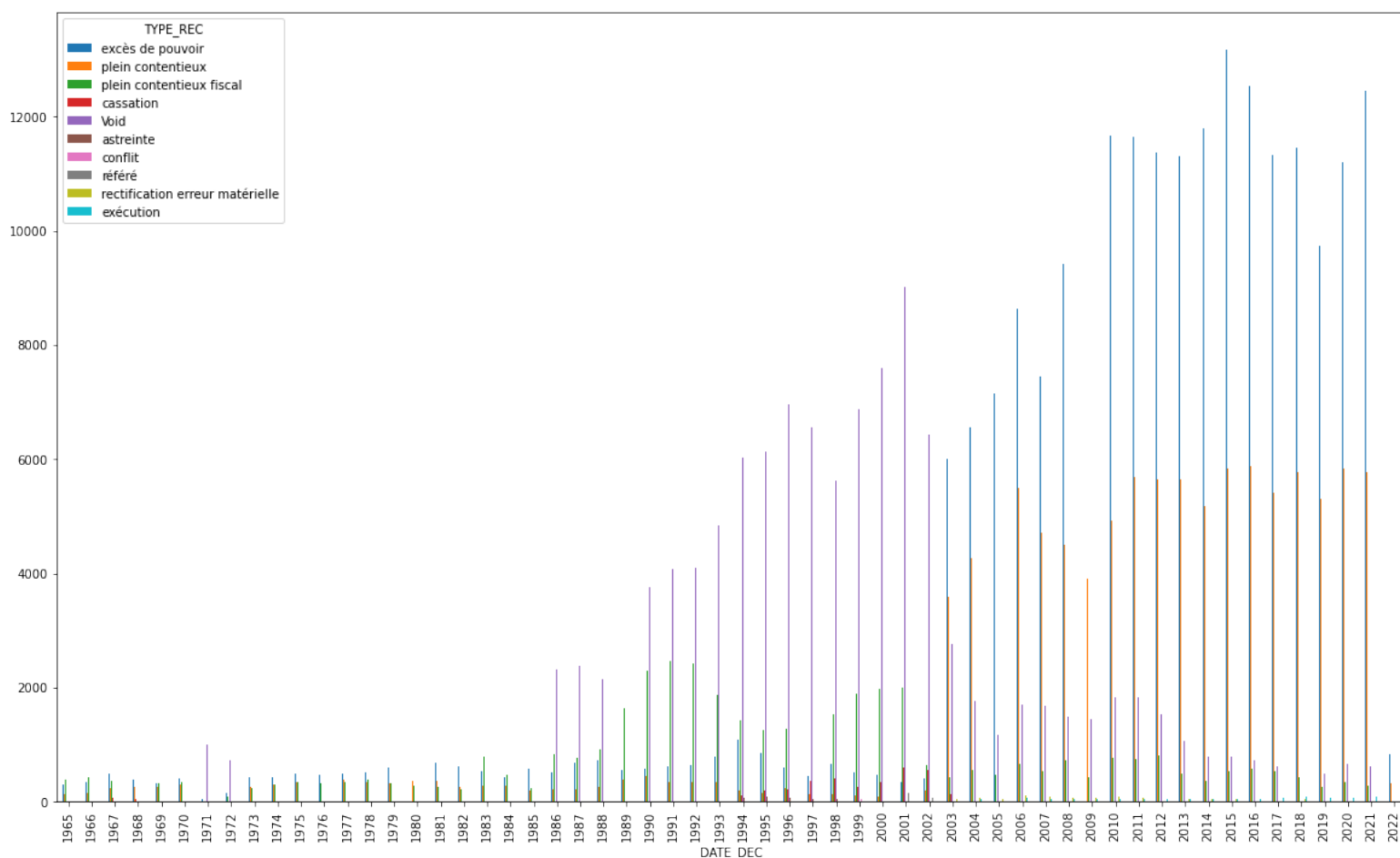
1 sur 2

19/05/2022 11:25

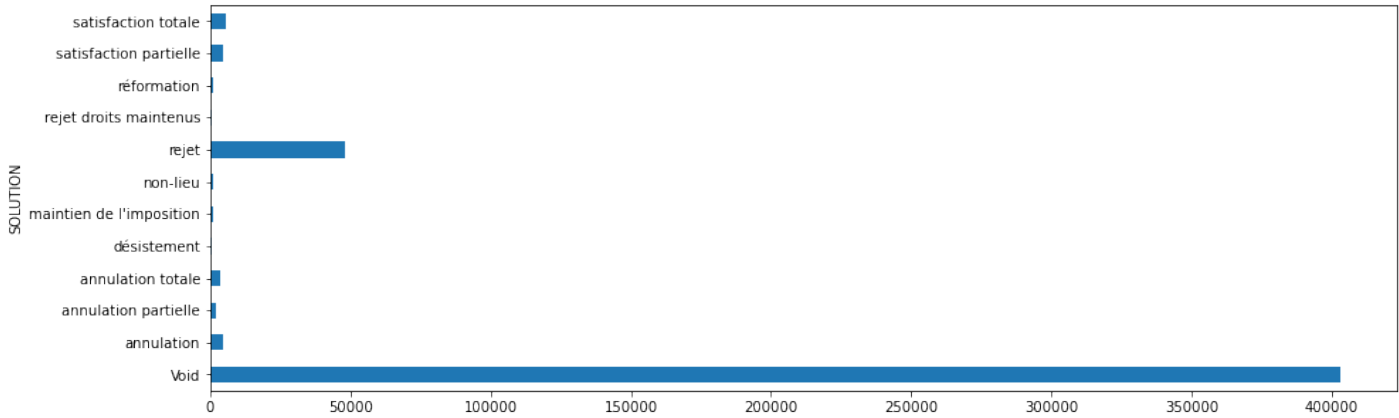
4. Graphique des types de recours qui sont mentionnés plus de 50 fois au sein des décisions de justice



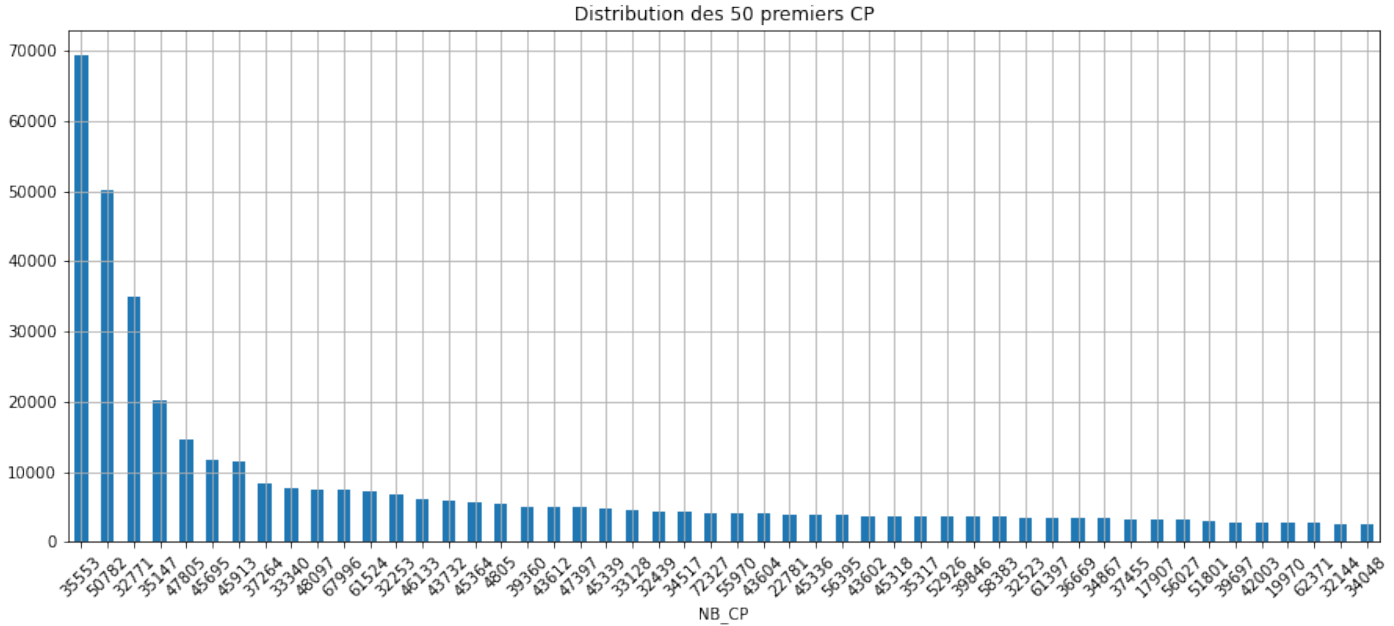
5. Graphique des types de recours par année sur les recours mentionnés plus de 50 fois au sein des décisions de justice



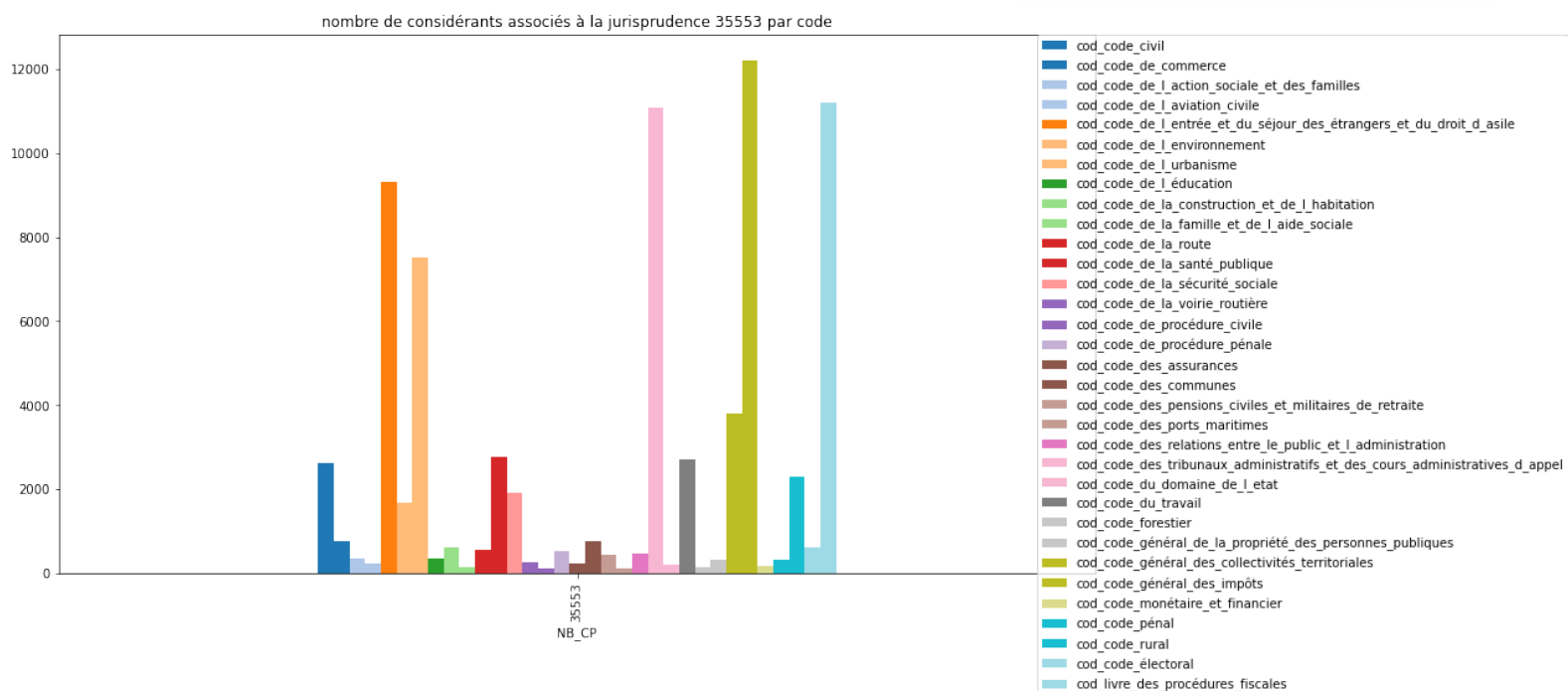
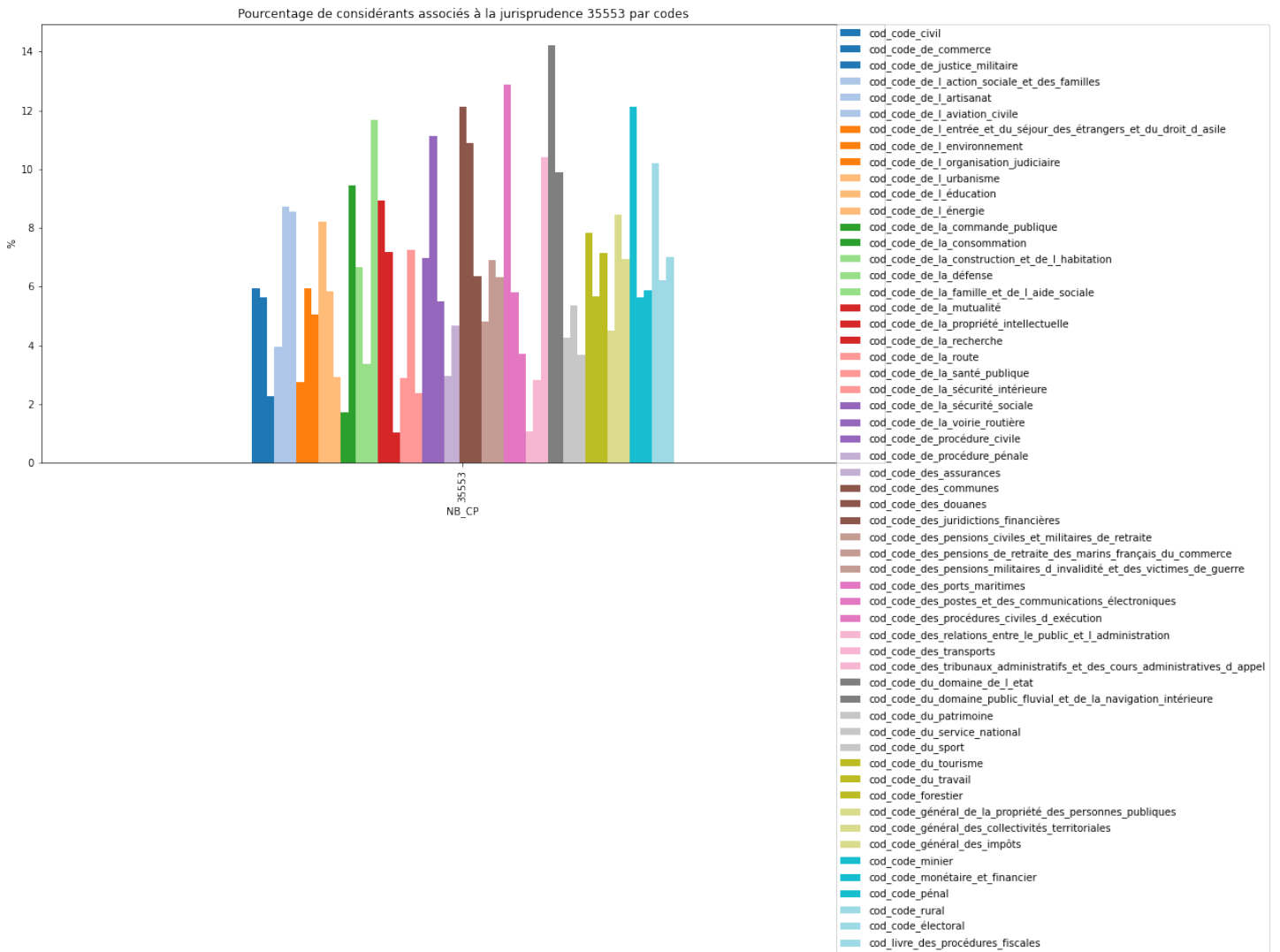
6. Graphique des solutions



7. Graphique des 50 jurisprudences les plus utilisées

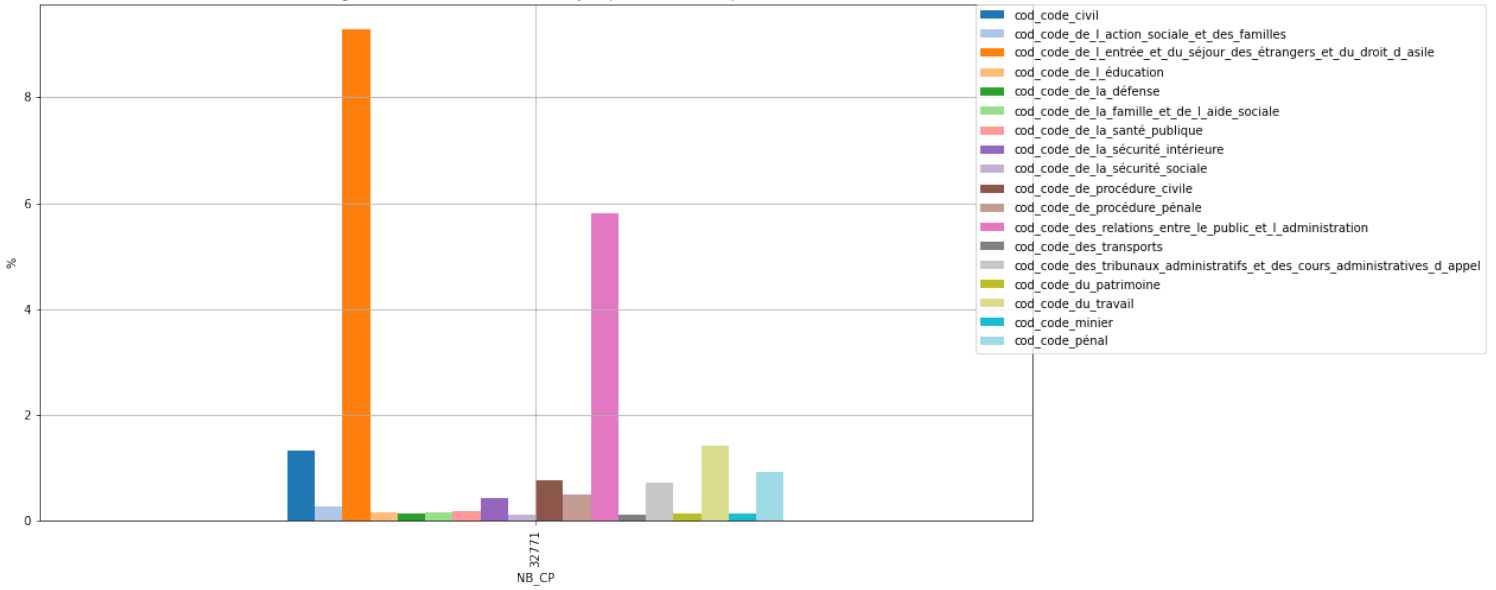


8. Graphiques des références légales de la jurisprudence 35553

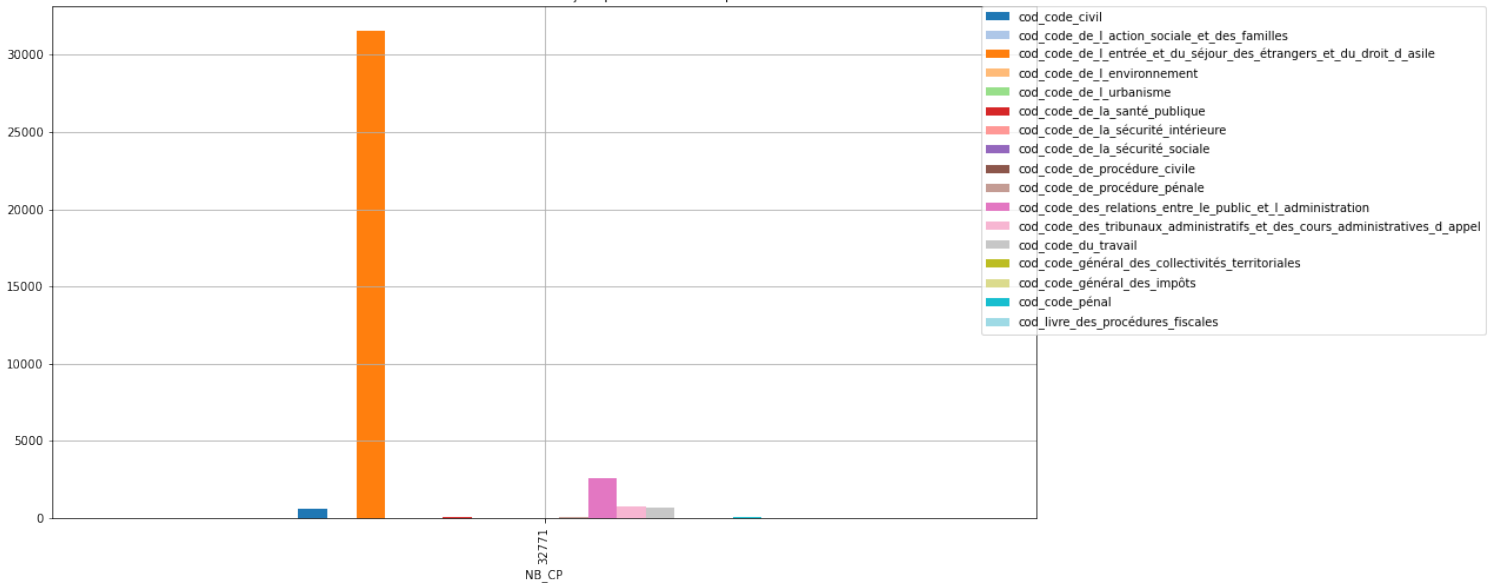


9. Graphiques des références légales de la jurisprudence 32771

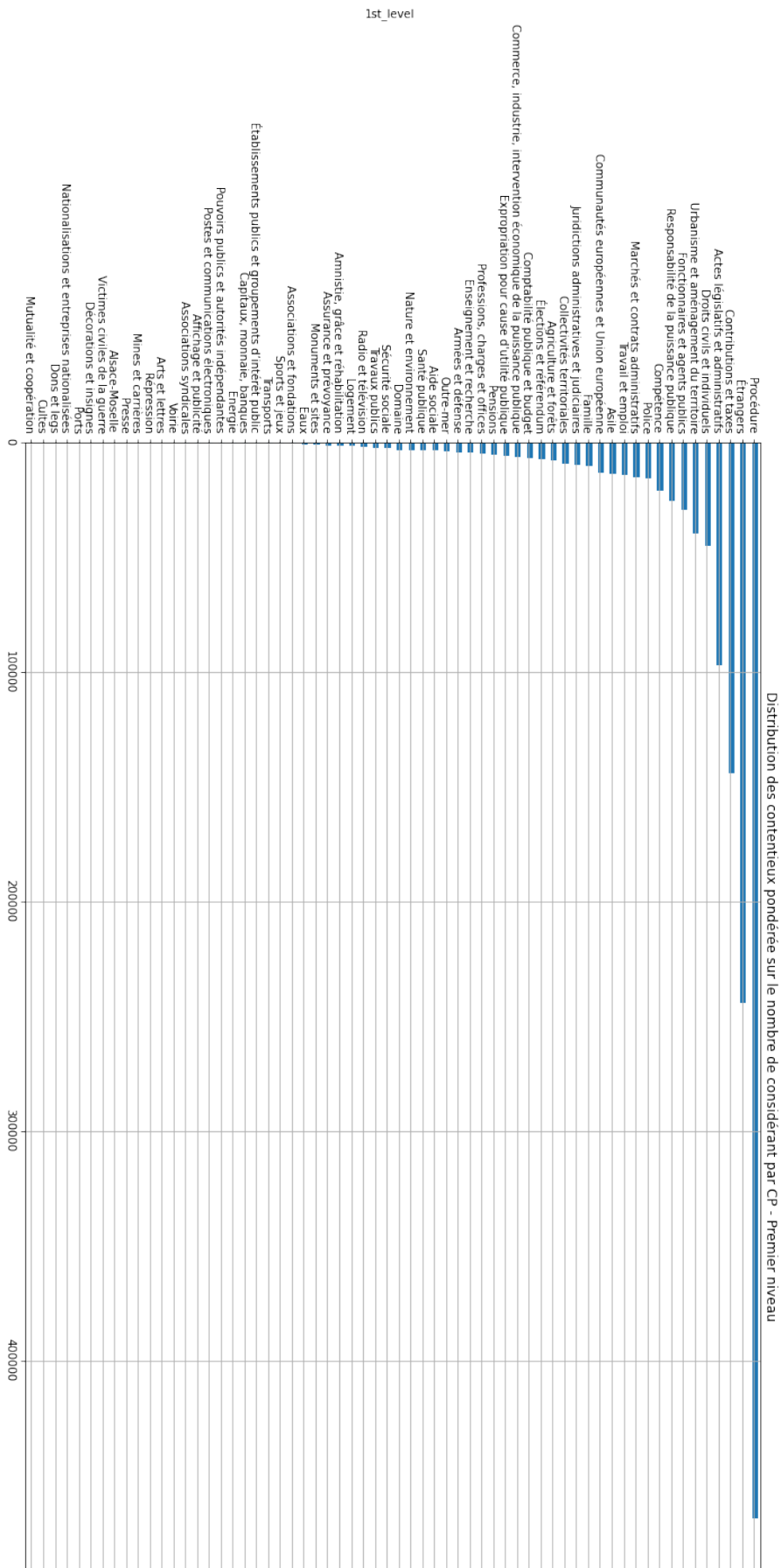
Pourcentage de considérants associés à la jurisprudence 32771 par codes



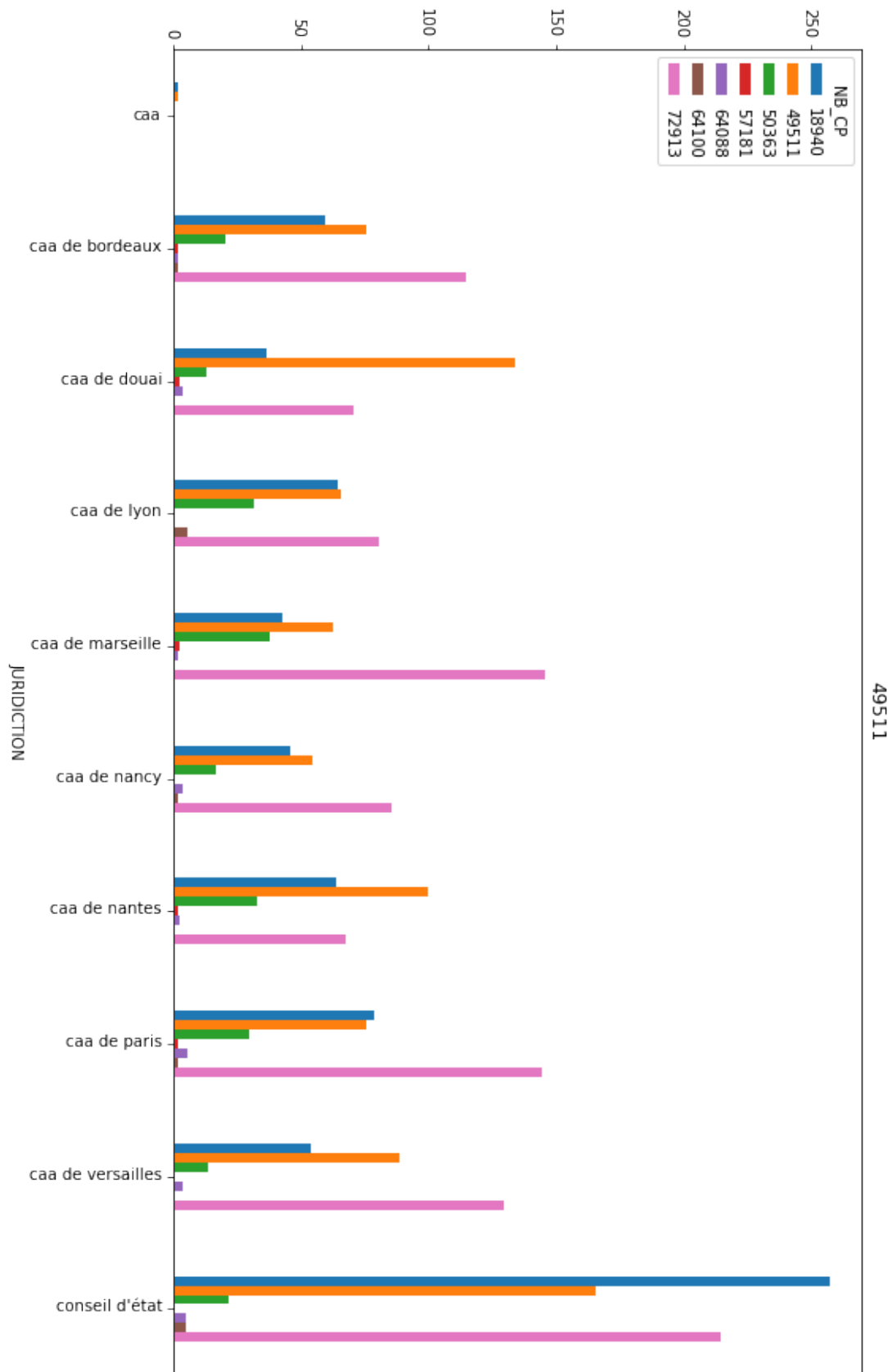
nombre de considérants associés à la jurisprudence 32771 par code



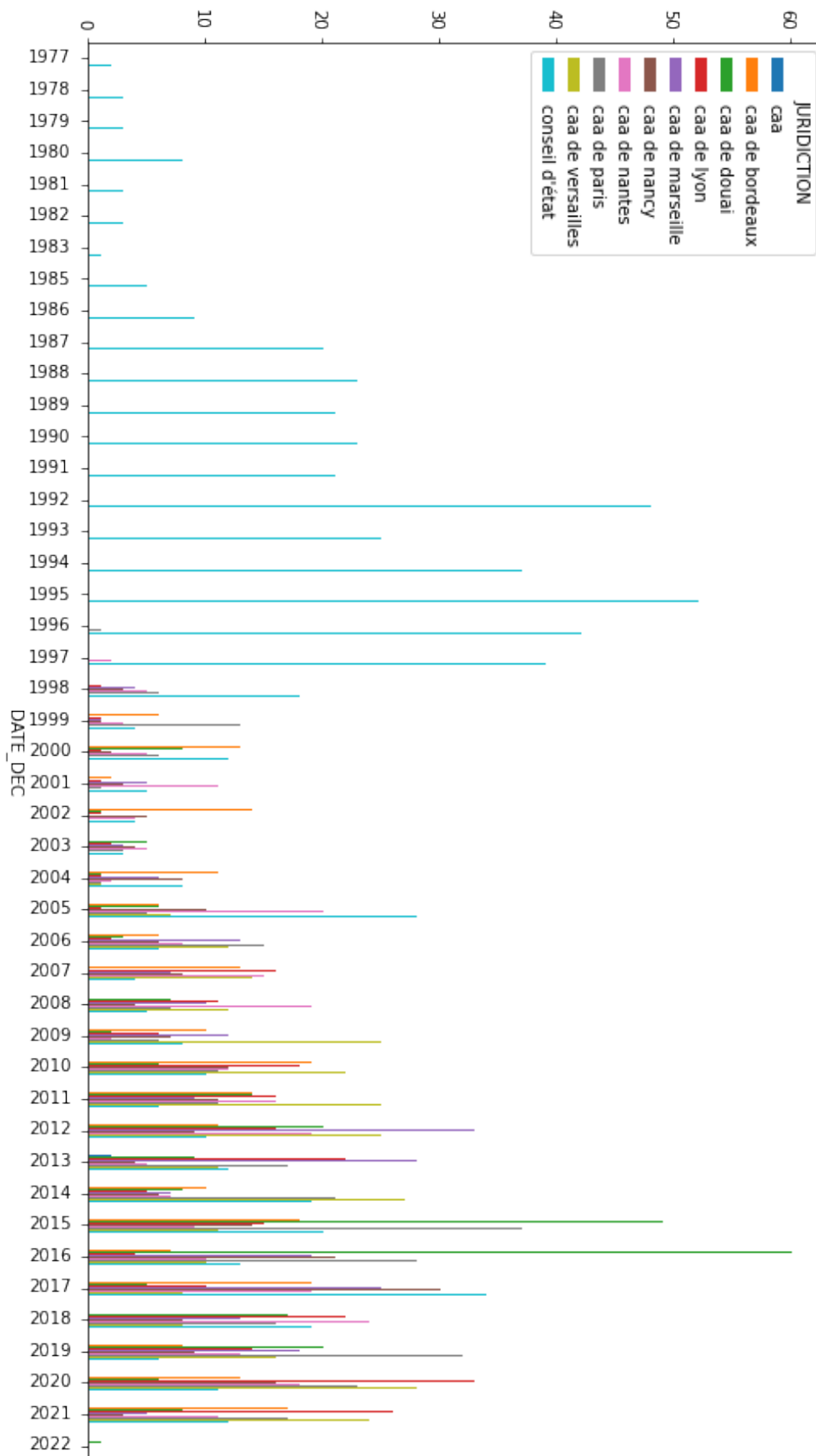
10. Distribution des contentieux pondéré sur le nombre de considérant par jurisprudence.



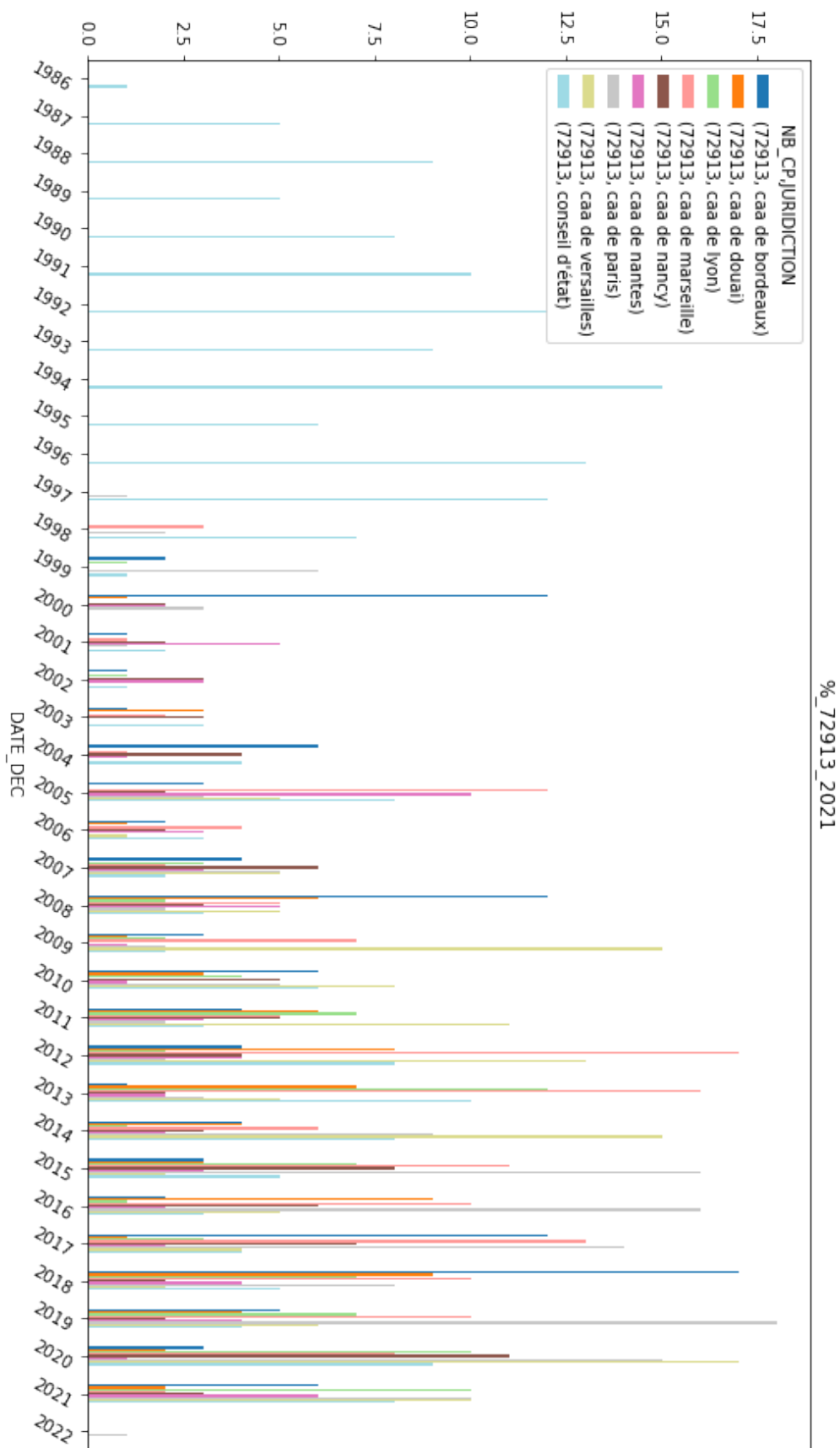
11. Distribution des considérants associés à la jurisprudence 49511, et jurisprudence proches, en fonction de la juridiction



12. Distribution de l'ensemble des jurisprudences proche de 49511, en fonction de l'année et de la juridiction



13. Distribution des considérants par année et par juridiction pour la jurisprudence 72913 en pourcentage par année



2. Exemples de jurisprudences

1. Exemples des 7 jurisprudences les plus mentionnées

35553 : « 54-05-03, 54-06-05-11 L'auteur d'une intervention n'étant pas partie à l'instance, les dispositions de l'article 75-I de la loi du 10 juillet 1991 font obstacle à ce que la partie perdante soit condamnée à payer la somme qu'il demande au titre des frais exposés et non compris dans les dépens. »

50782 : « 54-06-05-11 Les dispositions de l'article L. 761-1 du code de justice administrative font obstacle à ce que soit mise à la charge d'une partie à l'instance la somme qu'une personne qui n'a été appelée en la cause que pour produire des observations et n'a pas à ce titre acquis la qualité de partie demande au titre des frais exposés par elle et non compris dans les dépens. »

32771 : « 01-01-02-01-01, 335-02-04, 335-02-09, 54-07-01-04-035 Un étranger peut utilement se prévaloir, à l'appui de conclusions tendant à l'annulation de la mesure d'expulsion dont il a fait l'objet, des stipulations de l'article 8 de la Convention européenne de sauvegarde des Droits de l'homme et des libertés fondamentales aux termes desquelles : « 1°) Toute personne a Droit au respect de sa vie privée et familiale, de son domicile et de sa correspondance - 2°) Il ne peut y avoir ingérence d'une autorité publique dans l'exercice de ce Droit que pour autant que cette ingérence est prévue par la loi et qu'elle constitue une mesure qui, dans une société démocratique, est nécessaire à la sécurité nationale, à la sûreté publique, au bien-être économique du pays, à la défense de l'ordre et à la prévention des infractions pénales, à la protection de la santé ou de la morale, ou à la protection des Droits et libertés d'autrui ». »

35147 : « 54-06-05-11(1) Le décret n° 88-907 du 2 septembre 1988 ayant été abrogé par le décret n° 91-1266 du 19 septembre 1991 portant application de la loi n° 91-647 du 10 juillet 1991, les conclusions de la société Libertés Services S.A. doivent être regardées comme demandant la condamnation de l'Etat sur le fondement de l'article 75-8 de ladite loi ; Aux termes du 8 de l'article 75 de la loi du 10 juillet 1991 : "Dans toutes les instances le juge condamne la partie tenue aux dépens ou à défaut, la partie perdante, à payer à l'autre partie la somme qu'il détermine, aux frais exposés et non compris dans les dépens. Le juge tient compte de l'équité ou de la situation économique de la partie condamnée. Il peut, même d'office, pour des raisons tirées des mêmes considérations, dire qu'il n'y a pas lieu à cette condamnation ; Il n'y a pas lieu, dans les circonstances de l'espèce, de faire application des dispositions susmentionnées de la loi du 10 juillet 1991. »

47805 : « 335-03-03 Il résulte de l'ensemble des dispositions des articles L. 776-1 et L. 776-2 du code de justice administrative que les conclusions incidentes du requérant tendant à ce que la responsabilité pour faute de l'Etat soit engagée et à ce que l'Etat soit condamné à lui verser une indemnité ne sont pas au nombre de celles dont le président

du tribunal administratif ou le magistrat délégué par celui-ci, statuant sur le fondement des articles L. 776-1, L. 776-2, R. 776-1 et suivants du code de justice administrative, peuvent connaître. De telles conclusions ne peuvent faire l'objet que d'une demande présentée devant le juge administratif de Droit commun, statuant selon la procédure ordinaire. Ainsi, en jugeant qu'il y avait lieu de rejeter les conclusions incidentes présentées par le requérant tendant à la condamnation de l'Etat à lui payer des dommages et intérêts au motif qu'aucune faute ne pouvait être imputée à l'autorité administrative, le magistrat délégué par le président du tribunal administratif a méconnu l'étendue des pouvoirs qu'il tient des articles L. 776-1, L. 776-2, R. 776-1 et suivants du code de justice administrative. Dès lors, il y a lieu d'annuler le jugement de ce magistrat en tant qu'il a rejeté les conclusions indemnitaires présentées par le requérant et de renvoyer l'affaire sur ce point au tribunal administratif. »

45695 : « 335-01-02-02-01 Aux termes du 3° de l'article 12 bis de l'ordonnance du 2 novembre 1945 : "Sauf si sa présence constitue une menace pour l'ordre public, la carte de séjour temporaire portant la mention "vie privée et familiale" est délivrée de plein Droit : (...) A l'étranger, ne vivant pas en état de polygamie, qui justifie par tout moyen résider en France habituellement depuis plus de dix ans ou plus de quinze ans, si, au cours de cette période, il a séjourné en qualité d'étudiant". Il résulte de ces dispositions, éclairées par les travaux préparatoires de la loi n° 98-349 du 11 mai 1998, que l'étranger qui, dans les dix premières années suivant son entrée en France, a demandé et obtenu un titre de séjour en qualité d'étudiant, ne peut prétendre à la délivrance de plein Droit d'une carte de séjour temporaire portant la mention "vie privée et familiale" qu'après avoir résidé habituellement en France pendant plus de quinze ans. »

45913 : « 335-01-03-02, 335-03-01-01 Aux termes de l'article 12 bis de l'ordonnance du 2 novembre 1945, dans sa rédaction issue de la loi du 11 mai 1998 : "Sauf si sa présence constitue une menace pour l'ordre public, la carte de séjour temporaire portant la mention "vie privée et familiale" est délivrée de plein Droit : (...) 11° A l'étranger résidant habituellement en France dont l'état de santé nécessite une prise en charge médicale dont le défaut pourrait entraîner pour lui des conséquences d'une exceptionnelle gravité, sous réserve qu'il ne puisse effectivement bénéficier d'un traitement approprié dans le pays dont il est originaire". Aux termes de l'article 7-5 introduit dans le décret du 30 juin 1946 par le décret du 5 mai 1999 : "Pour l'application du 11° de l'article 12 bis de l'ordonnance du 2 novembre 1945 précitée, le préfet délivre la carte de séjour temporaire, au vu de l'avis émis par le médecin inspecteur de santé publique de la direction départementale des affaires sanitaires et sociales compétente au regard du lieu de résidence de l'intéressé. A Paris, l'avis est émis par le médecin, chef du service médical de la préfecture de police. Cet avis est émis dans les conditions fixées par arrêté du ministre chargé de l'intégration, du ministre chargé de la santé et du ministre de l'intérieur (...)". L'arrêté du 8 juillet 1999 pris pour l'application de ces dispositions impose au médecin inspecteur de santé publique de la direction départementale des affaires sanitaires et sociales d'émettre un avis précisant si l'état de santé de l'étranger nécessite ou non une prise en charge médicale, si le défaut de cette prise en charge peut ou non entraîner des conséquences d'une exceptionnelle gravité et si l'intéressé peut ou non bénéficier effectivement d'un traitement médical approprié dans son pays. Il appartient ainsi au médecin inspecteur, tout en respectant le secret médical, de donner au préfet les éléments relatifs à la gravité de la pathologie présentée par l'étranger intéressé et à la nature des traitements qu'il doit suivre, nécessaires pour éclairer sa décision. En se fondant sur un avis rendu par le médecin chef du service

médical de la préfecture de police qui se limitait à la seule mention manuscrite "séjour non justifié", la décision de refus de séjour et par voie de conséquence, l'arrêté de reconduite à la frontière ont été pris suivant une procédure irrégulière et sont par suite entachés d'illégalité. »

2. Jurisprudences 49511 et « proches »

18940 ("03-02-01[1], 66-07-01[2] Les salariés légalement investis de fonctions représentatives bénéficient, dans l'intérêt de l'ensemble des travailleurs qu'ils représentent, d'une protection exceptionnelle prévue par la loi. Lorsque le licenciement d'un de ces salariés est envisagé, ce licenciement ne doit pas être en rapport avec les fonctions représentatives normalement exercées ou l'appartenance syndicale de l'intéressé. Dans le cas où la demande de licenciement est motivée par un comportement fautif, il appartient à l'inspecteur du travail ou à l'inspecteur des lois sociales en agriculture saisi et, le cas échéant, au ministre compétent de rechercher si les faits reprochés au salarié sont d'une gravité suffisante pour justifier son licenciement compte tenu de l'ensemble des règles applicables au contrat de travail de l'intéressé et des exigences propres à l'exécution normale du mandat dont il est investi ; en outre, pour refuser l'autorisation sollicitée, l'autorité administrative a la faculté de retenir des motifs d'intérêt général relevant de son pouvoir d'appréciation de l'opportunité, sous réserve qu'une atteinte excessive ne soit pas portée à l'un ou l'autre des intérêts en présence.", 'A') PCJA : ['Agriculture et forêts']

49511 (" 66-07-01-04 En vertu des dispositions du code du travail, le licenciement des salariés légalement investis de fonctions représentatives, qui bénéficient d'une protection exceptionnelle dans l'intérêt de l'ensemble des travailleurs qu'ils représentent, ne peut intervenir que sur autorisation de l'inspecteur du travail. Lorsque leur licenciement est envisagé, celui-ci ne doit pas être en rapport avec les fonctions représentatives normalement exercées ou avec leur appartenance syndicale. Dans le cas où la demande de licenciement est motivée par un acte ou un comportement du salarié survenu en dehors de l'exécution de son contrat de travail, notamment dans le cadre de l'exercice de ses fonctions représentatives, il appartient à l'inspecteur du travail, et le cas échéant au ministre, de rechercher, sous le contrôle du juge de l'excès de pouvoir, si les faits en cause sont établis et de nature, compte tenu de leur répercussion sur le fonctionnement de l'entreprise, à rendre impossible le maintien du salarié dans l'entreprise, eu égard à la nature des fonctions de l'intéressé et à l'ensemble des règles applicables à son contrat de travail.", 'A') PCJA : ['Travail et emploi']

50363 (" 54-08-02-02-01-03 En vertu du code du travail, les salariés protégés bénéficient, dans l'intérêt de l'ensemble des travailleurs qu'ils représentent, d'une protection exceptionnelle. Lorsque le licenciement d'un de ces salariés est envisagé, il ne doit pas être en rapport avec les fonctions représentatives normalement exercées par l'intéressé ou avec son appartenance syndicale. Dans le cas où la demande de licenciement est motivée par l'inaptitude physique consécutive à un accident du travail, il appartient à l'inspecteur du travail, et, le cas échéant, au ministre, de rechercher, sous le contrôle du juge, si cette inaptitude est telle qu'elle justifie le licenciement du salarié, compte tenu des caractéristiques de l'emploi exercé à la date à laquelle elle est constatée, de l'ensemble des règles applicables au contrat de travail de l'intéressé, des exigences propres à l'exécution normale du mandat dont il est investi, et de la possibilité d'assurer son reclassement dans l'entreprise selon les modalités et conditions définies par l'article

L. 122-32- 5 du code du travail. En jugeant que l'inaptitude d'un salarié est de nature à justifier son licenciement, une cour administrative d'appel se livre à une appréciation souveraine des faits de l'espèce qui n'est pas susceptible d'être discutée devant le juge de cassation.", 'B') PCJA : ['Procédure']

57181 (" 66-07-01-04 1) Un agissement du salarié intervenu en dehors de l'exécution de son contrat de travail ne peut motiver un licenciement pour faute, sauf s'il traduit la méconnaissance par l'intéressé d'une obligation découlant de ce contrat.,,2) Dans le cas où la demande de licenciement est motivée par un acte ou un comportement du salarié qui, ne méconnaissant pas les obligations découlant pour lui de son contrat de travail, ne constitue pas une faute, il appartient à l'inspecteur du travail, et le cas échéant au ministre, de rechercher, sous le contrôle du juge de l'excès de pouvoir, si les faits en cause sont établis et de nature, compte tenu de leur répercussion sur le fonctionnement de l'entreprise, à rendre impossible le maintien du salarié dans celle-ci, eu égard à la nature de ses fonctions et à l'ensemble des règles applicables au contrat de travail de l'intéressé.", 'A') PCJA : ['Travail et emploi']

64088 (" 66-07-01-04-02 1) Lorsque le licenciement d'un salarié légalement investi de fonctions représentatives est envisagé, celui-ci ne doit pas être en rapport avec les fonctions représentatives normalement exercées ou avec son appartenance syndicale. Dans le cas où la demande de licenciement est motivée par un comportement fautif, il appartient à l'inspecteur du travail saisi et, le cas échéant, au ministre compétent, de rechercher, sous le contrôle du juge de l'excès de pouvoir, si les faits reprochés au salarié sont d'une gravité suffisante pour justifier le licenciement, compte tenu de l'ensemble des règles applicables au contrat de travail de l'intéressé et des exigences propres à l'exécution normale du mandat dont il est investi. Un agissement du salarié intervenu en dehors de l'exécution de son contrat de travail ne peut motiver un licenciement pour faute, sauf s'il traduit la méconnaissance par l'intéressé d'une obligation découlant de ce contrat.,,2) En l'espèce, salarié protégé ayant, lors d'une suspension de séance du comité d'établissement, asséné un violent coup de tête à un autre salarié. Cet acte de violence délibérément commis sur la personne d'un collègue sur le lieu du travail, même à l'occasion des fonctions représentatives de l'intéressé, doit être regardé comme une méconnaissance par celui-ci de son obligation, découlant de son contrat de travail, de ne pas porter atteinte, dans l'enceinte de l'entreprise, à la sécurité d'autres membres du personnel.", 'B') PCJA : ['Travail et emploi']

64100 (" 66-07-01-04-02 1) Lorsque le licenciement d'un salarié légalement investi de fonctions représentatives est envisagé, celui-ci ne doit pas être en rapport avec les fonctions représentatives normalement exercées ou avec son appartenance syndicale. Dans le cas où la demande de licenciement est motivée par un comportement fautif, il appartient à l'inspecteur du travail saisi et, le cas échéant, au ministre compétent, de rechercher, sous le contrôle du juge de l'excès de pouvoir, si les faits reprochés au salarié sont d'une gravité suffisante pour justifier le licenciement, compte tenu de l'ensemble des règles applicables au contrat de travail de l'intéressé et des exigences propres à l'exécution normale du mandat dont il est investi. Un agissement du salarié intervenu en dehors de l'exécution de son contrat de travail ne peut motiver un licenciement pour faute, sauf s'il traduit la méconnaissance par l'intéressé d'une obligation découlant de ce contrat.,,2) En l'espèce, salarié protégé ayant utilisé 105 heures de délégation pour exercer une activité salariée au sein d'une autre entreprise. L'utilisation par un salarié protégé de ses heures de délégation pour exercer une autre activité professionnelle méconnaît l'obligation de loyauté à l'égard de son employeur qui découle de son contrat de travail.", 'B') PCJA : ['Travail et emploi']

72913 (" 54-04-04 En vertu des dispositions du code du travail, les salariés légalement investis de fonctions représentatives qui bénéficient, dans l'intérêt des travailleurs qu'ils représentent, d'une protection exceptionnelle, ne peuvent être licenciés qu'avec l'autorisation de l'inspecteur du travail. Lorsque le licenciement d'un de ces salariés est envisagé, il ne doit pas être en rapport avec les fonctions représentatives normalement exercées par l'intéressé ou avec son appartenance syndicale.Dans le cas où la demande est motivée par un comportement fautif, il appartient à l'inspecteur du travail saisi et, le cas échéant, au ministre compétent, de rechercher, sous le contrôle du juge de l'excès de pouvoir, si les faits reprochés au salarié sont d'une gravité suffisante pour justifier son licenciement, compte tenu de l'ensemble des règles applicables à son contrat de travail et des exigences propres à l'exécution normale du mandat dont il est investi.....Lorsqu'il résulte d'un procès-verbal de constat d'huissier de justice, effectué en application l'article 1er de l'ordonnance n° 45-2592 du 2 novembre 1945, que les faits reprochés au salarié protégé sont établis, ces constatations font foi jusqu'à preuve contraire. Dès lors, il ne saurait être retenu que, en confrontant ces constatations à des attestations de salariés qui ne rapportent pas la preuve contraire, un doute subsiste qui doit profiter au salarié.", 'B') PCJA : ['Procédure']