CANDIDATURE POUR LA JOURNÉE DE LA JEUNE RECHERCHE EN PROPRIÉTÉ INTELLECTUELLE (7ÈME ÉDITION)

1. COORDONNÉES

Prénom, nom : Evgeniia VOLKOVA
Sujet de thèse : La gouvernance des données pour les systèmes d'IA de mobilité urbaine
Directeur de thèse : Jessica EYNARD, Maître de conférences en droit privé à l'Université Toulouse 1 Capitole
Université de rattachement : École Doctorale Droit et Science Politique de l'Université Toulouse 1 Capitole / ANITI
Première inscription en thèse : Décembre 2020 (3ème année de thèse)

2. SUJET DE THÈSE

L'objectif de cette recherche est une réflexion sur un cadre légal et éthique pour la gouvernance des données en matière de mobilité urbaine intelligente. Les villes modernes utilisent de plus en plus les systèmes d'intelligence artificielle (IA) afin de rendre la mobilité urbaine plus sûre, propre et adaptée aux besoins des citoyens. Par exemple, certaines villes européennes, comme Toulouse, ont expérimenté la navette autonome de niveau 4 qui est basée sur l'IA. D'autres exemples d'usage de l'IA sont la gestion du trafic, les feux de circulation intelligents ou l'utilisation des algorithmes d'apprentissage automatique pour définir les zones à risque d'accident.

Le développement et déploiement des systèmes d'IA nécessitent une disponibilité large des données de mobilité et des données connexes de toutes les parties prenantes de l'écosystème local pour créer des algorithmes et ensuite appliquer ces algorithmes aux données pour obtenir un résultat concret. Cependant, la qualité des jeux de données utilisés pour entraîner, tester et valider les systèmes d'IA de mobilité urbaine est tout aussi importante afin d'assurer que les résultants produits par ces systèmes soient corrects, non discriminatoires et sans biais. La disponibilité d'une grande quantité des données de mobilité urbaine de haute qualité permet de garantir que les caractéristiques culturelles, économiques et sociales, et les besoins de la population locale soient correctement représentés. Ces données permettent aux autorités locales de mettre en œuvre des solutions d'IA basées sur les données qui correspondent aux besoins des citoyens.

Or, les données de mobilité utilisées pour développer et mettre en œuvre les systèmes d'IA peuvent provenir de différentes sources et peuvent être à la fois personnelles et non personnelles. Cette nature complexe des données de mobilité les soumet à différents régimes juridiques de protection, ce qui rend leur gouvernance extrêmement difficile et peut compromettre la mise en œuvre d'un accès équitable aux données de mobilité et leur partage.

En particulier, l'appréhension progressive des données comme un objet de valeur pouvant être mis sur le marché et échangé contre une rémunération soulève des questions quant à la protection de ces données par les droits *in rem*, tels que le droit de propriété et droit de propriété intellectuelle (PI). Or, l'applicabilité de ces régimes de protection a récemment suscité de nombreux débats, tant au niveau politique et législatif européen et français, que dans la doctrine juridique.

Certains chercheurs affirment que le droit de propriété permettrait d'identifier les titulaires de droits[1] ainsi que de répartir de manière équitable des bénéfices économiques et des pouvoirs d'accorder des

1 Thomas HOEREN, « Big Data and the Ownership in Data: Recent Developments », *European Intellectual Property Review*, 2014, pp. 753-754.

droits d'utilisation sous licence[2]. En 2017, la Commission européenne a même proposé d'introduire le droit du producteur de données afin de faciliter l'accès aux données produites par les machines[3]. Or, la nature de ce droit n'est pas claire. Certains juristes le qualifient de nouveau type de droit de propriété intellectuelle[4] car le droit est structuré non pas comme un droit de propriété classique, comprenant le droit de posséder, d'utiliser et de disposer de la chose, mais comme un droit exclusif d'utiliser des données, y compris l'octroi de licences et le recours en justice en cas d'utilisation non autorisée[5]. La durée de ce droit n'a pas été précisée par la Commission, ce qui rapproche ce droit d'un droit de propriété absolu.

Or, le modèle de propriété est basé sur le principe du *numerus clausus* selon lequel les droits de propriété doivent être limités par des règles impératives en termes de nombre et de contenu, y compris les types d'objets de ces droits[6]. Bien que des exemples de reconnaissance indirecte existent en droit et dans la jurisprudence, aucun pays n'accorde des droits de propriété sur les données directement. Ainsi, si le « vol de données informatiques » a déjà été retenu par la jurisprudence française[7], cette jurisprudence ne peut être considérée comme une attribution de droit de propriété sur les données. Dans les pays de tradition civiliste, qui ne reconnaissent pas la règle du précédent, un nouveau droit *in rem* ne peut être créé par un tribunal.

En outre, plusieurs acteurs peuvent participer à la production des données de mobilité à différentes étapes de leur chaîne de valeur, tels que les fabricants de produits, les fournisseurs de services numériques et les utilisateurs, ce qui ne permet pas de définir un seul titulaire du droit sur les données coproduites[8]. En supposant que le droit du producteur de données soit attribué aux fabricants de produits, c'est *à priori* l'acteur privé qui doit être considéré comme un titulaire. Même si la proposition de la Commission prévoit que les relations entre plusieurs producteurs peuvent être régies par contrat, il existe un risque que le propriétaire privé de l'infrastructure impose des conditions inéquitables à l'utilisateur de données.

La possibilité de protéger les données par les droits de PI existants doit également être analysée. Par exemple, le droit d'auteur ne s'applique qu'à la forme d'expression, tandis que le contenu et les idées restent dans le domaine public[9]. Les données de mobilité doivent-elles également être exemptées de la protection du droit d'auteur sur la base de cet argument ? Aussi, les données de mobilité sont généralement produites automatiquement par des capteurs et des objets connectés. Dans quelle mesure cette caractéristique empêche la protection des données de mobilité par le droit d'auteur qui est basé sur l'exigence de « qualité de l'auteur » et de « créativité »[10] ?

---

2 Herbert ZECH, « Information as Property », *JIPITEC*, 2015, vol. 6, n° 3, p. 197.

3 Communication de la Commission du 10 janvier 2017 intitulée « Créer une économie européenne fondée sur les données », COM(2017) 9 final, p. 12.

4 Voir, par ex. : Bernt HUGENHOLTZ, *op. cit.*, p. 62. Voir supra n° 228 ; Dev S. GANGJEE, « The Data Producer's Right : An Instructive Obituary » (version préliminaire), *in* Ernest LIM et Phillip MORGAN (dir.), *The Cambridge Handbook of Private Law and Artificial Intelligence*, Cambridge : Cambridge University Press, 2022, p. 9. Disponible sur : https://ssrn.com/abstract=4051831 (consulté le 19/06/2023).

5 EUROPEAN COMISSION, Commission staff working document on the free flow of data and emerging issues of the European data economy Accompanying the Communication de la Commission du 10 janvier 2017, Brussels : SWD(2017) 2 final, 10 janvier 2017, p. 33.

6 Sjef VAN ERP, « Ownership of data: The numerus clausus of legal objects », *Brigham-Kanner Property Rights Conference Journal*, 2017, vol. 6, pp. 235 et 239.

7 Cass. crim., 20 mai 2015, n° 14-81336, PB.

8 Federico FERRETTI, « A Single European Data Space And Data Act For The Digital Single Market : On Datafication And The Viability Of Apsd2-Like Access Regime For The Platform Economy », *European Journal of Legal Studies*, août 2022, vol. 14, n° 1, p. 184.

9 Francesco BANTERLE, « Data Ownership in the Data Economy: A European Dilemma », *in* Tatiana-Eleni SYNODINOU et al. (dir.), *EU Internet Law in the Digital Era: Regulation and Enforcement*, 1e éd., Allemagne : Springer, 2019, p. 206.

10 Bernt HUGENHOLTZ, « Against 'data property' », *in* Hanns ULLRICH, Peter DRAHOS et Gustavo GHIDINI, *Kritika. Essays on Intellectual Property*, vol. 3, 2018, p. 55.

En outre, la directive 96/9/CE accorde la protection par le droit d'auteur à une bases de données[11]. Or, au moment de l'entrée en vigueur de la directive 96/9/CE, l'état de l'art connaissait des bases de données dites « SQL » ou « relationnelles »[12]. Ce texte serait-il applicable en ce qui concerne les données de mobilité actuellement collectées et stockées, soit dans des bases de données non relationnelles (NoSQL), soit en masse non structurée ?

De plus, la directive 96/9/CE n'applique que le critère d'originalité[13] et accorde le droit d'auteur aux bases de données représentant une « création intellectuelle propre à leur auteur » en raison « du choix ou de la disposition des matières »[14]. Dans l'affaire *Football Dataco Ltd*, la CJUE a développé le critère d'originalité d'une base de données indiquant qu'elle doit exprimer la « capacité créative de l'auteur » qui fait ses « choix libres et créatifs » et, donc, impliquer la « touche personnelle » de l'auteur. Le critère d'originalité ne sera pas rempli tant que la base de données n'est constituée qu'en fonction de considérations techniques qui n'impliquent pas une « liberté créative ».[15] Cela dit, la plupart des bases de données modernes sont créées sur la base des considérations techniques et répondent donc difficilement aux critères d'originalité établis par la CJUE.

La protection des bases de données de mobilité par le droit *sui generis* est également douteuse. Selon la doctrine, les moyens consacrés à la création des capteurs pour collecter des données ou l'analyse des données ne doivent pas être pris en compte pour évaluer le caractère substantiel de l'investissement dans la constitution de base de données, parce que l'objectif de ces activités, qui peuvent nécessiter des investissements importants, est de créer des données et d'inclure ces données dans la base de données[16]. D'autres auteurs font une distinction entre la création des données *stricto sensu* (par exemple, les données proviennent de la seule source et le fabricant produit des données dans le cadre d'une autre activité principale) et l'obtention de données existantes par collecte ou mesure via des capteurs[17]. En conséquence, les investissements dans la constitution d'infrastructures dédiées à la mesure ou à l'obtention de données existantes seront pris en compte lors de l'évaluation du caractère substantiel. La décision du législateur d'exempter les bases de données produites par les machines du champ d'application du droit *sui generis* dans le texte de la proposition de règlement sur les données du 23 février 2022[18] n'a fait que renforcer ces débats.

Par ailleurs, le secret des affaires est considéré comme un instrument efficace de protection des données, car il permet d'empêcher leur appropriation illicite par des tiers[19]. Alors qu'il n'accorde pas de droits exclusifs, il prévoit des recours juridiques contre des tiers en cas d'acquisition, d'utilisation et de divulgation illicites d'un secret protégé. Les données de mobilité peuvent être potentiellement considérées comme un objet du secret des affaires. Or, ces données sont difficilement conformes aux exigences du secret des affaires. En particulier, ces données sont souvent collectées par des capteurs

---

11 Directive 96/9/CE, art. 1 (2).

12 Daniel GERVAIS, « Exploring the Interfaces Between Big Data and Intellectual Property Law », *JIPITEC*, 2019, vol. 10, n° 1, p. 8.

13 Directive 96/9/CE, cons. 16.

14 *Ibid.*, art. 3, para. 1.

15 CJUE, 1 mars 2012, n° C-604/10, *Football Dataco Ltd e.a. c. Yahoo! UK Ltd e.a.*, para. 38-42.

16 Josef DREXL, « Designing Competitive Markets for Industrial Data : Between Propertisation and Access », *JIPITEC*, 2017, vol. 8, n° 4, p. 268.

17 Matthias LEISTNER, « Big Data and the EU Database Directive 96/9/EC: Current Law and Potential for Reform » *in* Sebastian LOHSSE, Reiner SCHULZE et Dirk STAUDENMAYER (dir.), *Trading Data in the Digital Economy: Legal Concepts and Tools : Münster Colloquia on EU Law and the Digital Economy III*, Allemagne : Nomos, 2017, p. 28.

18 Proposition de règlement du Parlement européen et du Conseil du 23 février 2022 fixant des règles harmonisées pour l'équité de l'accès aux données et de l'utilisation des données (règlement sur les données), COM/2022/68 final.

19 Tanya APLIN, Alfred RADAUER, Martin A. BADER et al., « The Role of EU Trade Secrets Law in the Data Economy : An Empirical Analysis », *IIC – International Review of Intellectual Property and Competition Law* [en ligne], 2023, https://doi.org/10.1007/s40319-023-01325-8.

dans les espaces publics et ne peuvent donc pas constituer un secret. De plus, même si les données peuvent avoir une valeur commerciale potentielle, le lien avec le caractère secret des données est difficile à prouver. En outre, la manière dont les données sont collectées et partagées rend difficile le maintien de leur caractère secret, même si certaines mesures de protection sont mises en place. Par conséquent, la protection des données de mobilité en tant que secret des affaires n'est pas évidente et doit être évaluée au cas par cas.

Ainsi, l'objectif est d'évaluer de manière critique les implications des régimes juridiques existants, y compris le droit de propriété, les droits de PI et le secret des affaires, pour la disponibilité et la qualité des données de mobilité et le développement et déploiement des systèmes d'IA de mobilité urbaine dignes de confiance et respectueux des intérêts et droits des détenteurs de données privées.

3. AXES DE RECHERCHE SPÉCIFIQUES

1) Quel est l'impact de la protection des données de mobilité par les droits de PI sur la mise en œuvre des systèmes d'IA dignes de confiance et acceptables ?

Les systèmes d'IA modernes basés sur l'apprentissage (automatique et en profondeur) sont souvent décrits comme « gourmands en données », ce qui signifie que leur développement nécessite une grande quantité de données. L'exactitude, la précision, la fonctionnalité et l'efficacité de ces systèmes dépendent directement de la quantité et de la qualité des données utilisées pour les entraîner, les tester et les valider. Cependant, les incertitudes juridiques existant quant à la protection de ces données par les droits de PI peuvent avoir un impact important sur la quantité et la qualité des données disponibles sur le marché.

Quel est le rôle du droit pour garantir un accès équitable et sans restriction à ces données sur le marché européen ? Les données de mobilité doivent-elles être protégées sur la base des droits de PI existants, y compris le droit d'auteur et le droit *sui generis*, ou sur la base du secret des affaires ? Quel est l'impact de l'appréhension des données comme un objet de propriété / une marchandise et la possibilité de les protéger par des droits exclusifs, y compris le droit de propriété (intellectuelle) ? L'approche doit-elle être différente pour les données à caractère personnel et non personnel ?

Le législateur devrait-il plutôt introduire un nouveau droit exclusif de protection des données ? Quels étendue et critères de protection devraient être définis ? Qui doit être considéré comme titulaire de droits (fabricant de l'appareil / objet connecté, fournisseur de services numériques, utilisateur final) ? Quels sont les modèles alternatifs de gouvernance des données ? Les données doivent-elles plutôt être traitées comme un bien public exempt de tout régime de protection exclusif et librement accessible à toute personne souhaitant développer des systèmes d'IA ?

Dans quels cas l'accès obligatoire aux données peut-il être imposé sans entraver de manière significative les droits de PI des détenteurs de données et l'innovation ? Des exceptions devraient-elles être prévues, par exemple, pour permettre le développement de systèmes d'IA qui servent l'intérêt public, tels que l'amélioration des services publics, la mobilité urbaine et la santé publique, et pour lutter contre le changement climatique et la propagation des maladies ?

2) Quelles implications de la protection des données par le secret des affaires pour les systèmes d'IA de mobilité urbaine ?

Le secret des affaires est souvent présenté comme un instrument flexible de protection des systèmes d'IA et des données sous-jacentes qui, associé à une protection contractuelle, peut permettre aux titulaires de droits d'accorder ou de restreindre l'accès à ces systèmes et à ces données. Or, la protection des données de mobilité par le secret des affaires entraîne plusieurs incertitudes juridiques.

La première difficulté consiste à identifier le détenteur du secret des affaires, que la directive 2016/943 ne définit que de manière vague, comme toute personne physique ou morale qui contrôle un secret de façon licite. Cette définition a été transposée en droit français en s'appuyant également sur la notion de contrôle licite de données. Pourtant, ni la directive européenne, ni la loi française la transposant ne clarifient les notions de « contrôle licite » ni de personne concrète exerçant ce contrôle. Cette notion de contrôle laisse une marge large d'interprétation quant à savoir qui contrôle les données de mobilité produites en utilisant des capteurs, appareils mobiles ou autres IoT : le fabriquant d'un objet connecté qui contrôle les données *de facto* ou l'utilisateur de cet objet qui le détient physiquement et l'exploite ?

De plus, par rapport aux droits de PI et, en particulier, au droit *sui generis* qui permettent la fouille de données, le cadre légal ne prévoit pas de telles exceptions pour le secret des affaires. D'une part, le secret d'affaires ne peut être assimilé aux droits de PI dans le sens qu'il ne confère au détenteur légitime aucun droit absolu. D'autre part, et en revanche, la protection des données par le secret des affaires peut mettre en péril les exceptions légales.

En outre, même si le secret des affaires sert à protéger les informations confidentielles des entreprises contre leur appropriation illicite par les concurrents, il empêche également l'accès à toute information sur les algorithmes et les explications derrière la prise de décision automatisée, ce qui entrave la transparence du système d'IA[20]. Le fournisseur privé d'un système d'IA peut demander à l'autorité publique de ne pas divulguer son algorithme et ses documents liés au développement et déploiement en affirmant qu'il s'agit de secrets des affaires[21].

Quel est l'impact de ces restrictions d'accès aux données de mobilité sur l'opacité des algorithmes ? Lorsque les données et les algorithmes sont gardés secrets, il y a un risque de les soustraire à l'examen public et au contrôle démocratique. La situation est particulièrement délicate dans le cas où des algorithmes privés protégés par le secret des affaires sont utilisés par les autorités publiques dans la prise des décisions, comme le système d'IA américain COMPAS. La loi doit-elle imposer un accès obligatoire dans un tel cas pour protéger l'intérêt public ?

En dernier lieu, un système d'IA développé à partir de données de mobilité protégées par le secret des affaires, qui ont été appropriées et utilisées de manière illicite, doit-il être considéré comme un bien en infraction ? Les autorités judiciaires peuvent interdire la production et la commercialisation des biens en infraction ou ordonner leur retrait ou destruction. La comparaison avec d'autres juridictions, serait-elle pertinente dans le contexte européen ? Par exemple, l'art. 2 (1)(x) de la loi japonaise de 2015 sur la prévention de la concurrence déloyale limite la portée des biens en infraction aux choses créées par des actes d'utilisation non autorisée des secrets techniques (à savoir des informations techniques protégées par le secret des affaires), ce qui est défini de manière très étroite par rapport à la directive 2016/943.

20 Katarina FOSS-SOLBREKK, « Three routes to protecting AI systems and their algorithms under IP law : The good, the bad and the ugly », *Journal of Intellectual Property Law & Practice*, 2021, vol. 16, n° 3, p. 258.

21 Robert BRAUNEIS et Ellen P. GOODMAN, « Algorithmic Transparency for the Smart City », *Yale Journal of Law & Technology*, 2018, vol. 20, p. 153.