



Trouver le diamant

dans la mine de données ou les implications juridiques de l'exploration de données

[recherche] La conjonction du big data et du data mining offre de nombreuses perspectives aux chercheurs. Pour profiter pleinement de ces promesses, il est urgent de trouver des solutions juridiques.

Gâce à l'accroissement constant de la capacité de stockage, d'analyse et de transmission de données, la technologie numérique a fait de tels progrès ces dernières années que les chercheurs scientifiques peuvent maintenant faire appel à des techniques de recherche qui étaient auparavant hors de leur portée. Ces techniques de recherche, dites d'exploration de données, ne sont pas nécessairement nouvelles, mais elles sont devenues aujourd'hui de plus en plus accessibles et intéressantes dans tous les domaines

scientifiques confondus. La fouille dans des bases de données, qu'elles contiennent des données brutes (résultats d'expériences, de mesures ou d'enquêtes), des textes, des images ou tout autre type d'œuvres, permet de réaliser des liens et des rapprochements entre différents éléments d'information, là où il était autrefois impossible de le faire. Cette exploration numérique des données laisse miroiter non seulement des bénéfices économiques importants, en terme d'économie de temps et d'effort, donc d'argent, mais également en terme de

découvertes attendues, dans le domaine des sciences médicales et environnementales notamment.

Le fouillage, le minage ou l'exploration de données

Il n'existe pas encore de définition universellement acceptée de l'exploration de données mais, de façon générale, il s'agit de méthodes de recherche ayant pour objet l'extraction d'un savoir ou d'une connaissance à partir de grandes quantités de données, grâce à des procédés automatiques ou semi-automatiques. Les sources consultées par les ////

//// chercheurs sont aussi multiples que diverses : sites ouverts de réseaux sociaux (tels Facebook et Twitter), référentiels de données gérés soit par des centres de recherche, des universités ou tout autre type d'institution publique, ou sites ouverts ou payants maintenus par des entreprises privées, incluant des éditeurs scientifiques (tels Elsevier, LexisNexis et Wiley). Un accès facile et à des conditions raisonnables aux sources de données est par conséquent essentiel afin de pouvoir effectuer les recherches qui mèneront aux progrès scientifiques si souhaités. Or, l'accès et la réutilisation de données à des fins de recherche font souvent face à des obstacles de nature juridique : les détenteurs de bases de données soit en restreignent l'accès par l'application de mesures techniques, soit en limitent la réutilisation par des tiers, se réclamant de leurs droits de propriété intellectuelle sur ces bases.

La protection des bases de données par la propriété intellectuelle

Qu'en est-il de la protection de bases de données et de leur contenu en droit français et européen ?

Si les articles et les monographies publiés par les chercheurs sont généralement protégés par la propriété littéraire et artistique en tant qu'œuvres originales de l'esprit, il n'en va pas nécessairement de même des bases de données ou de leur contenu. Une base de données sera jugée suffisamment originale en droit d'auteur si la sélection et l'arrangement de son contenu dénotent un niveau suffisant d'originalité pour rencontrer les critères de protection du Code de la propriété intellectuelle. Or, une base de données scientifique ne manifeste généralement pas le niveau d'originalité requis par le droit d'auteur. En science, en effet, on vise davantage à son exactitude et à son exhaustivité plutôt qu'à son caractère créatif. En revanche, la constitution, la vérification ou la présentation du contenu d'une base de données donnent lieu en France à

un droit exclusif distinct, au titre de l'article L.341-1 du Code de la propriété intellectuelle, si cette base de données atteste d'un investissement financier, matériel ou humain substantiel. Cette protection est indépendante et s'exerce sans préjudice de celles résultant du droit d'auteur ou d'un autre droit sur la base de données ou l'un de ses éléments constitutifs. C'est dire que si la base de données réunit des œuvres originales, le producteur jouira d'un droit distinct sur la base, en plus des droits pouvant persister sur les œuvres qu'elle contient.

Quels sont les critères de protection de ce droit sur les bases de données ?

La protection juridique des bases de données prévue dans le Code français découle de la directive européenne 1996/9/EC du 11 mars 1996. Cette directive introduisit à l'époque un régime de protection inédit dans le monde, dont les contours ont dû être plus amplement définis par les tribunaux nationaux et par la Cour de justice européenne. Il ressort de la jurisprudence de la Cour de Luxembourg que les frais encourus pour la création de données incluses dans une base de données ne sont pas pris en compte dans le calcul de l'investissement substantiel. Le droit de protection est consenti en récompense à la collecte, la vérification et la présentation de données existantes, et non à la création de nouvelles données¹. Une protection a été rejetée pour les données découlant de l'organisation de matchs de football et de courses hippiques, mais a été acceptée pour les données recueillies pour la publication de catalogues de vente ou de répertoires d'annonces. La question de savoir si une base de données rencontre les critères de protection est une question qui, de fait, se décide au cas par cas.

Quelle est l'étendue de la protection conférée au titre du droit des bases de données ?

Le titulaire d'un droit sur une base de données jouit du droit exclusif d'autoriser l'extraction et la réuti-

lisation d'une partie substantielle du contenu de sa base. Ces deux droits reçoivent généralement une interprétation large de la part de la Cour européenne. Dans une décision récente, la Cour a jugé que la recherche dans une base de données grâce à un métamoteur de recherche dédié constitue un acte de contrefaçon².

Or, l'activité d'exploration d'un métamoteur de recherche n'est pas éloignée d'un acte d'exploration de données tel qu'entendu dans le domaine scientifique.

Et l'exception pour fins de recherche ?

La directive 1996/9/EC prévoit certes une disposition permettant aux États membres d'introduire une exception pour fins d'enseignement et de recherche mais, comme cette disposition est facultative, elle a été transposée de différentes manières à travers l'Europe. Pour sa part, le législateur français a transposé cette disposition d'une manière très restrictive à l'article L.342-3, 4, laissant ainsi très peu de place aux activités d'exploration de données à des fins de recherche.

Il n'est donc pas exclu qu'un chercheur s'adonnant dans le cadre de ses recherches à l'exploration de données enfreigne les droits du titulaire d'une base de données protégée au sens de l'article L. 341-1 du Code de la propriété intellectuelle.

Cet état de chose crée d'ores et déjà une situation d'incertitude au sein de la communauté scientifique, qui se manifeste par une hésitation généralisée à faire usage des outils numériques en recherche.

Les solutions envisageables

L'importance de l'exploration de données croît au même rythme que les développements de la technologie numérique.

Si l'Europe en général, et la France en particulier, désirent demeurer compétitives au niveau mondial sur le plan de la recherche scientifique, une solution devra être apportée au problème naissant.



Concilier abonnements électroniques et commande publique

Trois solutions sont possibles :

- une réforme du droit des bases de données, de façon à exempter les actes d'exploration pour fins de recherche du champ du droit exclusif ;
- l'accroissement de l'utilisation de licences libres pour la mise à disposition de bases de données résultant de recherches financées au moyen de fonds publics ;
- la négociation de licences avec les titulaires de droits tels que les éditeurs scientifiques de manière à permettre l'exploration de données contenues dans leurs bases à des conditions acceptables.

Cette dernière solution constituait d'ailleurs le principal cheval de bataille de la Commission européenne dans le cadre de l'initiative, lancée au courant de l'année 2013, intitulée « Licences for Europe »³. Ce dialogue structuré entre parties intéressées n'a malheureusement abouti à aucun consensus, que ce soit à propos de l'identification du problème ou des solutions à y apporter.

Une chose est claire cependant, à savoir que le problème de l'exploration des données ne saura être résolu que par une combinaison des solutions proposées ci-dessus. L'avenir de la recherche en Europe en dépend. ■

➤ Lucie Guibault

Institut du droit de l'information (IVIR),
Université d'Amsterdam

L.Guibault@uva.nl

1. Voir notamment quatre décisions de la CJCE du 9 novembre 2004 : The British Horseracing Board Ltd e.a. ; Fixtures Marketing Ltd c/ Organismos Pronostikon ; Fixtures Marketing Ltd c/Oy Veikkaus AB ; Fixtures Marketing Ltd c/ Svenska Spel AB

2. Voir CJCE du 19 décembre 2013 Innoweb/ Wegener.

3. http://ec.europa.eu/internal_market/copyright/licensing-europe/index_fr.htm

[abonnement] **Transparence, égalité et libre concurrence sont les principes qui réglementent les achats publics. Les contrats conclus pour fournir des ressources électroniques n'échappent pas à la règle. Comment alors ajuster les besoins particuliers d'abonnements numériques au cadre des marchés publics ? La question se pose parfois avec acuité.**

La variété et l'évolutivité des ressources électroniques, tant en ce qui concerne leurs contenus, leurs fonctionnalités, leurs types d'accès, leurs modèles économiques et tarifaires, leurs modes d'acquisition, que les acteurs concernés (éditeur/agrégateur/producteur/diffuseur) en font un produit complexe à acquérir. Les professionnels de la documentation doivent négocier¹ des flux et des « bibliothèques de licences » avant de gérer des accès, payables à l'avance malgré les risques de dysfonctionnement. La gestion de temporalités antagonistes entre abonnements, marchés et procédures administratives devrait être anticipée dans chaque clause faisant référence aux dates, durées ou délais, ainsi que dans la clause tarifaire².

Tendre vers une offre adaptée et facile d'accès

Les établissements ont pour objectif de mettre à disposition des usagers une offre numérique adaptée et facile d'accès. Pour ce faire, les besoins en documentation électronique sont évalués en fonction de différents critères : missions³, répartition géographique⁴, destination⁵, statut⁶, variation tarifaire⁷, modes d'accès et types d'authentification⁸, moyens d'exploitation disponibles⁹, types de signalement et d'exploitation des métadonnées¹⁰, actions de formation, de médiation, de retours d'usages, impératifs internes ou locaux¹¹. Ce recueil de besoins fait l'objet d'un cahier des charges précis et prospectif. Les prévisions des délais nécessaires aux procédures puis à la mise en œuvre de reprise de portefeuille sont planifiées. En réponse aux besoins, il existe une diversité d'offres qui ont toutes des spécificités, et qui ne sont généralement pas substituables. C'est pourquoi, selon nous, la mise en concurrence, lancée uniquement sur la base du seuil, ne devrait pas porter sur les ressources elles-mêmes mais sur leur souscription.

Dans le cadre légal¹² du Code des marchés publics (CMP), la dépense s'évalue par famille homogène, à définir par l'établissement. Selon les types de pouvoirs adjudicateurs, des seuils - révisés régulièrement - sont définis pour déclencher des types de procédures, avec ou sans publicité. Le législateur encourage l'allotissement pour définir des ensembles plus circonscrits de fournitures et ainsi mieux rencontrer les offres des candidats, selon un découpage laissé au libre choix du pouvoir adjudicateur. L'exclusivité est une exception prévue par la loi. Les ressources concernées sont alors souscrites directement *via* un marché négocié monté avec le fournisseur, à condition que celui-ci fournisse une attestation d'exclusivité.

Rechercher un équilibre entre les acteurs

Sauf exception, ce sont des agences d'abonnement, historiquement spécialisées dans les périodiques, qui répondent aux appels ///