

Producteur de base de données : Leboncoin.fr obtient partiellement gain de cause

Le TGI de Paris a reconnu qu'un site de petites annonces constitue une base de données et que son exploitant a la qualité de producteur de base de données. Mais son jugement du 1er septembre 2017 montre, une fois de plus, qu'obtenir la condamnation d'actes de reproduction et de réutilisation de données sur le fondement du droit sui generis du producteur n'est pas chose aisée. Si le site Leboncoin.fr a obtenu gain de cause sur le fondement de l'article L. 342-2 du CPI relatif à l'extraction ou la réutilisation répétée et systématique de parties non substantielles de sa base, il a cependant vu ses demandes rejetées au titre de l'article L. 342-1 concernant le droit d'interdire ces actes sur une partie quantitativement ou qualitativement substantielle de son contenu. Pour la gestion des utilisateurs mécontents de voir leurs annonces reproduites, Leboncoin.fr obtient 20 000 € de dommages-intérêts, 20 000 € pour son préjudice d'image et 10 000 € au titre des frais engagés pour la procédure.

Entreparticuliers.com propose aux particuliers depuis 2000 un service payant d'hébergement d'annonces essentiellement immobilières. Leboncoin.fr, qui a été ouvert en France en 2006, permet de diffuser gratuitement des annonces pour tous types de biens. La simplicité d'utilisation de cet outil et sa gratuité ont concouru à son très grand succès. Pour maintenir un volume d'annonces, Entrepaticuliers.com a souscrit auprès d'un sous-traitant un service de piges lui fournissant toutes les nouvelles annonces immobilières publiées par les particuliers en France, dont les annonces publiées sur Leboncoin.fr et reprises sur Entrepaticuliers.com sans autorisation. Le premier a assigné le second en contrefaçon sur le fondement du droit d'auteur et du droit sui generis du producteur.

Le tribunal a commencé par reconnaître que le site Leboncoin constitue bien une base de données protégée par le droit d'auteur et que la société qui l'exploite peut bénéficier de la protection du droit du producteur, dit droit sui generis. Il a reconnu que cette dernière avait consenti d'importants investissements pour la constitution, la vérification et la présentation des données dans sa base, condition pour bénéficier de cette protection. Mais l'exploitant n'a obtenu que très partiellement gain de cause sur les extractions de contenu. Au regard du nombre d'annonces reproduites, le tribunal a jugé que les données en cause ne constituaient pas une partie quantitativement substantielle de la base de données. Sur les 28 millions d'annonces postées sur Leboncoin.fr en effet, 800 000 correspondaient au secteur immobilier. Le tribunal a estimé par ailleurs que le demandeur ne démontrait pas davantage l'extraction qualitativement substantielle de sa base. En revanche, il a obtenu gain de cause sur le caractère répété et systématique d'extraction et de réutilisation des données, au titre de l'article L. 342-2. Ces pratiques excèdent manifestement les conditions d'utilisation normale de la base de données.

Affaire Ryanair / Opodo : open data, producteur de bases de données et investissement substantiel.

Par Antoine Cheron, Avocat.- mercredi 25 mars 2015

La Cour de cassation vient de rendre un arrêt très attendu des acteurs du commerce numérique et plus spécialement de ceux intervenant dans le domaine du tourisme. La Chambre commerciale s'est en effet prononcée sur le conflit qui oppose depuis 2010 la compagnie aérienne low cost Ryanair à l'agence de voyage en ligne la SAS Opodo (Chambre commerciale, 10 février 2015, Ryanair c/Opodo, n°12-26.023).

Dans un contexte amplement favorable à l'innovation numérique, à la circulation des données et à l'open data, beaucoup se sont étonnés des actions en justice menées par la compagnie Ryanair à travers l'Europe contre les agences de voyages en ligne qui reprennent les informations de la base de données Ryanair, relative aux vols, disponibilités, horaires et tarifs, pour proposer des billets d'avions et autres services.

Toutefois, la réaction de la compagnie Ryanair face à ce qu'elle estime être une extraction et une réutilisation substantielle du contenu de sa base de données constitue une défense classique, adoptée par tout titulaire de droits d'auteur subissant une atteinte à son droit d'exploitation.

En effet, l'article L.112-3 du CPI prévoit notamment que l'auteur d'une base de données peut jouir de la protection instituée par le droit d'auteur lorsque cette base, par le choix ou la disposition des matières, constitue une création intellectuelle. Et, pour mieux encore protéger l'investissement consacré au contenu d'une base de données, une protection additionnelle et indépendante est prévue à l'article L.341-1 du CPI, qualifiée de droit sui generis.

Lorsqu'ils subissent une atteinte à leurs droits portant sur des créations intellectuelles, les titulaires de ces droits revendiquent volontiers la protection juridique du droit d'auteur. Face à pareilles revendications, le juge veillera à rétablir l'auteur de l'œuvre dans ses droits. Mais ce faisant il doit s'assurer qu'il ne contribue pas à conférer un monopole ou la réservation d'un champ de création au bénéfice d'un titulaire, c'est-à-dire aboutir à une surprotection.

En effet, afin de favoriser la dynamique des créations intellectuelles il est nécessaire, paradoxalement, que la protection des droits d'auteur n'aboutisse pas à la réservation d'un marché empêchant toute nouvelle création.

Cela est particulièrement vrai avec l'environnement numérique où les créations utilitaires telles que les bases de données constituent une valeur essentielle de l'Internet et de la société de l'information, dont la circulation est à la base de l'innovation numérique.

Profitant de la protection juridique élevée que lui offre le droit d'auteur en matière de base de données, la compagnie Ryanair a dans cette affaire revendiqué cette protection, en invoquant le bénéfice du droit sui generis qui protège le producteur d'une base de données contre les extractions et réutilisations sans autorisation de sa base de données.

La condition posée par cet article afin de bénéficier de la protection par le droit sui generis est la réalisation par le producteur d'un investissement substantiel financier, matériel ou humain, dans la constitution, la vérification et la présentation de la base de données.

L'intérêt essentiel de la présente décision, en ce qui concerne la protection de la base de données, réside donc dans l'appréciation de l'investissement consacré à la constitution de la base de données. L'investissement réalisé était-il suffisamment substantiel sur le plan financier, matériel ou humain pour accorder à la base de données Ryanair la protection du droit sui generis ?

Faits et arguments en présence

La compagnie Ryanair qui souhaite commercialiser ses vols directement et sans intermédiaire sur son site Internet a constaté que la SAS Opodo, agence de voyage en ligne, proposait aux internautes visitant le site Opodo.fr de réserver des sièges sur les vols Ryanair, alors qu'aucune autorisation en ce sens ne lui avait été accordée.

Ryanair reproche à la SAS Opodo de scanner quotidiennement le contenu de son site et de réutiliser ce contenu sur le site Opodo.fr pour vendre les vols Ryanair, augmentés de frais et commissions pouvant représenter jusqu'à 30 % du billet, l'internaute étant tenu dans l'ignorance quant au bénéficiaire des commissions prélevées.

L'extraction et la réutilisation du contenu de la base de données Ryanair par Opodo porteraient ainsi préjudice à la compagnie aérienne, en violant les droits d'auteurs qu'elle détiendrait sur sa base de données. La compagnie soutenait également que l'utilisation de sa marque par la SAS Opodo caractérisait la contrefaçon de la marque Ryanair.

La compagnie Ryanair sollicitait donc la condamnation de la SAS Opodo à une somme avoisinant les 3 millions d'euros sur le fondement notamment de l'article L.341-1, L.342-1 et L.716-1 du CPI.

Pour sa part la SAS Opodo conteste à la compagnie Ryanair le bénéfice de la protection issue du droit sui generis prévu à l'article L.341-1 en soutenant une absence d'investissement substantiel pour l'obtention, la vérification et la présentation de la base de données Ryanair. Elle nie par ailleurs avoir effectué une extraction d'une partie qualitativement ou quantitativement substantielle du contenu de la base de données Ryanair qui est accessible au public.

Concernant l'argument de contrefaçon de la marque Ryanair, la société Opodo conteste avoir porté une atteinte à la fonction essentielle de la marque puisqu'elle utilise la marque de la compagnie pour commercialiser non pas ses propres services mais les vols Ryanair. De plus, l'agence de voyage soutient que l'article R.322-4 du Code de l'aviation lui impose sous peine de sanction de communiquer aux consommateurs l'identité du transporteur aérien, c'est-à-dire en l'espèce la compagnie Ryanair.

A titre reconventionnel la SAS Opodo sollicite la condamnation de la compagnie Ryanair à des dommages et intérêts pour avoir commis à son encontre un acte de concurrence déloyale en publiant dans la presse un article la dénigrant : « Opodo est tenu par un consortium de compagnies aériennes à tarifs élevés, Ryanair fait de la vente directe, les passagers devraient donc éviter les agents de voyage en ligne qui les arnaquent en leur imposant des surcharges à des prix exorbitants et trompent les passagers ».

Le TGI de Paris puis la Cour d'appel vont rejeter toutes les prétentions de la compagnie Ryanair et en particulier celle relative à la protection de sa base de données par le droit sui generis, bien que la cour ait reconnu parallèlement que « les informations et données constituent une base de données au sens de l'article L.112-3 du CPI, susceptible d'être protégée par le droit sui generis ».

La Cour condamnera par ailleurs la compagnie Ryanair à verser à la SAS Opodo 30.000 euros de dommages et intérêts au titre du dénigrement à son encontre.

Le pourvoi formé par la compagnie Ryanair reprochait à l'arrêt d'appel d'avoir notamment écarté la protection par le droit sui generis au motif qu'il n'est pas démontré que les investissements informatiques réalisés à hauteur de 462 000 euros mensuels l'ont été pour la constitution, la vérification et la présentation de la base de données.

Un second moyen viendra critiquer l'arrêt d'appel ayant retenu un usage licite de la marque Ryanair sur le site opodo.fr et soutiendra principalement que l'usage de la marque Ryanair par la société Opodo n'est pas conforme aux usages honnêtes en matière commerciale dès lors que cet usage vise à laisser penser qu'un lien commercial unit les deux sociétés.

La décision de la Chambre commerciale

La Chambre commerciale rejette le pourvoi. En ce qui concerne l'investissement, la Cour de cassation approuve la cour d'appel d'avoir « exclu du champ de l'investissement entrant dans la constitution, la vérification ou la présentation du contenu de ladite base, tant le coût des logiciels destinés à assurer le fonctionnement du système de gestion commerciale que les dépenses relatives à l'application informatique de la billetterie », lequel investissement ne revêtait donc pas un caractère substantiel.

A propos de l'usage de la marque Ryanair, la Cour retiendra que la SAS Opodo n'a fait usage du signe protégé « que pour désigner, de manière nécessaire, les services de transport aérien de Ryanair qu'elle proposait au consommateur ». La contrefaçon de marque n'est donc pas ici caractérisée.

L'absence d'investissement substantiel par le producteur de la base de données

Afin de garantir les intérêts économiques de la personne qui consacre des investissements pour collecter, vérifier et présenter des informations sous la forme d'une base de données, l'article L 341-1 du CPI lui offre une protection contre l'extraction et la réutilisation des données sans autorisation.

Il s'agit d'un droit sui generis instauré par la directive du 11 mars 1996 relative à la protection juridique des bases de données et indépendant selon la CJUE des règles prévues par le droit d'auteur [1]. Ainsi, le producteur d'une base de données peut bénéficier de la protection par le droit sui generis indépendamment de toute originalité des données ou de la base de données elle-même.

Pour bénéficier de cette protection il suffit d'apporter la preuve d'un investissement substantiel financier, matériel ou humain dans la constitution, la vérification et la présentation du contenu de la base de données.

En l'espèce, la compagnie Ryanair n'est pas parvenue à apporter la preuve d'un tel investissement dans la constitution du contenu de la base de données. Le fait que sa base de données ait été considérée comme originale au regard de l'article L.112-3 du CPI n'a pas d'incidence pour la mise en œuvre du droit sui generis.

Si la compagnie Ryanair a effectivement soumis au juge le rapport d'un expert informatique, attestant d'un investissement par la compagnie d'un montant de 15 millions d'euros sur quatre ans pour la mise en place et le fonctionnement de son système informatique de billetterie et logiciel de gestion commerciale, cet investissement financier et matériel ne concerne pas directement la constitution, la vérification et la présentation de sa base de données.

En d'autres termes, si la compagnie a échoué à apporter la preuve d'un investissement substantiel c'est en raison de la confusion qu'elle a opérée entre ses investissements techniques et informatiques et les moyens consacrés à la recherche des éléments existants, à leur vérification et à leur rassemblement au sein d'une base de données, selon les termes employés par l'arrêt de la CJCE du 9 novembre 2004 [2].

Le producteur d'une base de données qui revendique le bénéfice du droit sui generis doit selon la jurisprudence démontrer qu'il a consacré des moyens pour obtenir les données mais également pour en vérifier la fiabilité.

Or dans le domaine numérique la preuve de ces moyens est d'autant plus nécessaire que les acteurs de l'Internet ont la possibilité de collecter des données de manière automatisée et de vérifier leur exactitude à l'aide de logiciels. C'est pour cette raison que la jurisprudence, spécialement celle de la CJUE, est venue resserrer la notion d'investissement substantiel consacré à la constitution du contenu d'une base de données.

C'est ainsi que la Cour de cassation exige du producteur de la base de données qu'il rapporte la preuve de moyens consacrés cumulativement à la constitution et à la vérification des données. Ainsi, une société spécialisée dans les annonces immobilières n'a pu bénéficier de la protection du droit sui generis pour sa base de données car elle ne vérifiait pas la fiabilité des annonces proposées par les internautes [3].

En l'espèce, il y a lieu d'approuver la solution retenue par la Chambre commerciale dans la mesure où la compagnie Ryanair ne démontre à aucun moment qu'elle aurait consacré des moyens pour collecter les données et vérifier leur fiabilité ou leur mise à jour.

En savoir plus sur <https://www.village-justice.com/articles/Affaire-Ryanair-Opodo-open-data,19293.html#lpXAH44rfUFAFZdD.99>

Droit des données : comment scraper des données dans la légalité ?

Thomas Saint-Aubin, CEO Seraphin.legal et Charles Leconte, legal data scientist et co-fondateur de CaseIP

Si les données constituent l'un des principaux actifs immatériels des entreprises, la légalité du processus de collecte automatisée est un préalable indispensable pour pouvoir valoriser ultérieurement ce patrimoine informationnel.

Les savoir-faire en matière de capture, de stockage, d'analyse, de traitement, d'enrichissement et de visualisation des données sont particulièrement recherchés pour rester compétitif. Mais qu'en est-il de la légalité des données ? C'est l'émergence du legal data scientist.

Avant de pouvoir exploiter les données, il faut les collecter. Il existe différents moyens de collecter des datas :

1 soit manuellement, ce qui peut nécessiter un temps considérable lorsqu'on cherche à disposer d'un volume important ;

2 soit au moyen de méthodes automatiques, via des logiciels, ou scraper, permettant d'obtenir une quantité importante d'informations dans un laps de temps record.

De plus en plus de sociétés ont recours au web scraping pour récupérer le contenu des sites afin d'enrichir leur propre base ou de générer de nouveaux business.

La question de la légalité de cette activité est souvent traitée tardivement, notamment par les investisseurs dans le cadre de la « due diligence » (vérifications) ou encore au moment de la démarche de mise en conformité RGPD.

Mais comment le droit qualifie et encadre la collecte des données, particulièrement lorsqu'elle est opérée par des robots ? Quelles sont les précautions à prendre pour transformer cet actif numérique en actif valorisable ? Quel est l'état de l'art pour préfigurer un robot juriste scrapeur ?

Etat du droit applicable au scraping

A l'ère du big data et des pratiques généralisées du scraping des données, Me Nicolas Courtier remet en cause la protection sui generis des producteurs des bases de données de la loi de 1998. "Le droit des producteurs des bases de données repose sur une approche statique des traitements de données : on se concentre sur la création de la base et non sur son utilisation".

Du point de vue du propriétaire de la donnée scrapée, qui a investi du temps et de l'argent dans la construction d'une base afin de proposer des services à valeur ajoutée sur une plateforme web, la récupération du fruit de son travail est assimilable à du vol.

En droit positif, plusieurs textes permettent (droit des producteurs de bases de données, directive sur le secret d'affaires, loi Informatique et Libertés pour la protection des données personnelles, action en concurrence déloyale, etc.) permettent de poursuivre le scraping illégal.

En propriété intellectuelle, le droit sui generis consacré en 1998 (art. L342-1 du code de la propriété intellectuelle) permet au producteur de la base de données d'interdire, entre autre, « l'extraction, par transfert de la totalité ou d'une partie qualitativement ou quantitativement substantielle du contenu de la base de données sur un autre support, par tout moyen et sous toute forme que ce soit » ou encore « la réutilisation, par la mise à la disposition du public de la totalité ou d'une partie qualitativement ou quantitativement substantielle du contenu de la base, quelle qu'en soit la forme » (art. L342-1 du code de la propriété intellectuelle).

C'est sur ces fondements que le Tribunal de Grande Instance de Paris a condamné, le 1er septembre 2017, le site Entrepaticuliers.com (TGI de Paris, 1er septembre 2017, Leboncoin.fr c/ Entrepaticuliers.com).

En droit pénal, le législateur a fait de l'extraction de données un délit spécifique. La loi du 24 juillet 2015 a modifié l'article 323-3 du Code pénal qui réprime désormais le fait « d'extraire, de détenir, de reproduire, de transmettre » frauduleusement les données d'un Système de traitement automatisé de données (STAD). Le vol de donnée est donc bien un délit distinct du vol d'une chose matérielle.

La limite de l'application de cette règle au web scraping tient au fait que l'article ne mentionne uniquement les STAD et que le législateur n'en a retenu aucune définition et il semble que le juge en fasse une acceptation large. Ainsi la qualification a été retenue pour le réseau carte bancaire (TGI de Paris, 13eme Ch. Correctionnelle, 25 février 2000) ou encore un disque dur. Reste à savoir si un site web peut bénéficier de cette protection.

Comment empêcher le scraping ? Comment se protéger ?

Dans la majorité des cas, le site cible du scraping n'a pas intérêt à laisser le script accéder à ses données.

Aucune technique n'est infaillible mais ces méthodes permettent a minima d'associer à une protection juridique, une protection technique.

- La création d'un compte utilisateur (signature CGU)

L'une des techniques les plus simple pour éviter le scraping est d'exiger la création d'un compte afin de visualiser le contenu du site. Si la création d'un compte est exigée, l'administrateur peut tracer les actions des utilisateurs et facilement détecter les comportements suspects. Cependant, cette technique, qui peut-être dissuasive pour les bots, peut également l'être pour les utilisateurs...

- Le bannissement d'ip

Pour bloquer l'accès d'un site aux robots scrapers, l'une des premières actions qui peut être intenté est la vérification régulière des logs de connexion. En cas d'activité inhabituelle indiquant un accès automatisé, comme par exemple l'indication de plusieurs actions similaires à partir de la même adresse IP, il est toujours possible de bloquer ou limiter cet accès. En bannissant cette IP, le script est bloqué.

Toutefois, de nombreux systèmes utilisent des proxies permettant de changer régulièrement d'adresse IP. Alors la technique ne produit qu'un effet de ralentissement.

- Captchas

Il est possible de limiter le nombre d'action de l'utilisateur dans un temps déterminé. En cas de doute, il est possible d'avoir recours à des Captchas (« Completely Automated Test to Tell Computers and Humans Distinction ») permettant de valider l'identité de l'utilisateur. Cette méthode est également partiellement efficace.

Mais elle irrite les vrais utilisateurs et il existe des logiciels pour les contourner.

TAF : Votre employeur vous demande de constituer une base de données de prospects (futurs clients à contacter) pour permettre aux commerciaux de lancer une nouvelle campagne commerciale en contactant ces prospects.

Cette base de données sera constituée en « Scrapant » les données sur Linkedin (nom prénom profession entreprise ...).

Avant de réaliser le logiciel de scraping, vous devez identifier les règles juridiques liées à cette activité :-

- Définissez le web scraping
- De quelles protections juridiques peuvent bénéficier les BDD
- A votre avis, Linkedin peut-il bénéficier de ces protections ? Expliquez
- Quels sont ici les risques juridiques liés à cette demande ?
 - o RGPD
 - o Propriété intellectuelle
 - o Contrat
- Selon vous quel est le risque juridique principal en matière de scraping ?