

Décision n° 16-SOA-02 du 23 mai 2016
relative à une saisine d'office pour avis portant sur l'exploitation des
données dans le secteur de la publicité en ligne

L'Autorité de la concurrence,

Vu le traité sur le fonctionnement de l'Union européenne et notamment ses articles 101 et 102 ;

Vu le livre IV du code de commerce modifié et notamment son article L. 462-4 ;

Adopte la décision suivante :

I. Contexte

1. Dans sa communication sur l'économie des données, publiée en 2014, la Commission européenne rappelle que nous « *assistons à une nouvelle révolution industrielle portée par les données numériques, l'informatique et l'automatisation* » et que les « *activités humaines, les processus industriels et la recherche conduisent tous à la collecte et au traitement de données sur une échelle sans précédent, favorisant l'apparition de nouveaux produits et services ainsi que de nouveaux processus d'entreprise et méthodes scientifiques* »¹.
2. À cet égard, dans une étude parue en 2015, l'OCDE indique que selon certaines estimations, le cumul des données stockées aurait atteint en 2015 environ huit zettabytes, contre moins d'un zettabyte² en 2009. Ce chiffre pourrait être multiplié par quarante avant la fin de la décennie³.

¹ Communication de la Commission au Parlement européen, au Conseil, au Comité économique et social et au comité des régions, « Vers une économie de la donnée prospère », 2.7.2014, COM(2014)442 final.

² 1 zettabyte = 10²¹ bytes = 1000 exabytes.

³ OCDE, Data-driven innovation, Big Data for Growth and Well-Being, October 06, 2015.

3. Cette augmentation s'explique notamment par la diversification des sources de production de données : celles-ci sont en effet générées par les individus, mais aussi par les objets connectés, les sites web, les applications mobiles, les systèmes informatiques... Elle provient également de la plus grande diversité des données elles-mêmes. Par exemple, et sans que la liste ne soit exhaustive, les individus génèrent aujourd'hui plusieurs catégories de données : données de contact (nom, prénom, e-mail...), données sociodémographiques, données de transaction, données relationnelles, données comportementales (recherches, types d'achats, publicités et vidéos visionnées...), données de géolocalisation, données de communication (messagerie instantanée, e-mails...), données techniques (identifiants d'appareils)⁴. Et ce développement concerne au premier chef les données non structurées⁵ (fichiers audio, vidéos, fichiers textes, messages écrits, contenu des pages web...), qui représentent aujourd'hui la grande majorité des données produites⁶.
4. Afin d'exploiter ces nouvelles formes de données de façon automatisée et instantanée, des outils informatiques ont été développés au cours des années 2000, par de nouvelles entreprises telles qu'Amazon ou Google. Google a par exemple développé des infrastructures informatiques afin d'améliorer le fonctionnement de son moteur de recherche. Google est ainsi à l'origine du modèle de programmation MapReduce, qui a permis d'améliorer le fonctionnement du système d'indexation des pages web par son moteur de recherche et de pouvoir exploiter des volumes importants de données non structurées⁷. De manière comparable, afin de créer son système de recommandation automatisé à la fin des années 90, Amazon a développé des infrastructures informatiques pour ses propres besoins, puis a, dès 2006, commercialisé ces services à des entreprises tierces. Amazon Web Services est aujourd'hui le leader du secteur de l'informatique en nuage. Le marché mondial des services et des technologies d'exploitation des données massives (*Big Data*) devrait atteindre 48,6 milliards USD en 2019 contre 3,2 milliards USD en 2010, avec un taux de croissance d'environ 23 %⁸.
5. Ces dernières années, l'automatisation et l'exploitation de données se sont significativement accrues dans un des secteurs centraux de l'économie numérique, celui de la publicité en ligne. Depuis 2010, celui-ci a connu de fortes évolutions, parmi lesquelles le développement de l'achat programmatique, qui permet la « *mise en relation automatisée de vendeurs et d'acheteurs d'inventaires publicitaires* »⁹, et le développement des systèmes d'enchères en temps réel (ou RTB, *Real Time Bidding*), qui consistent à « *allouer une impression publicitaire à un client et d'en déterminer le prix en fonction des critères de ciblage et de ses caractéristiques comme la taille de la bannière, le contexte de la page web ou encore l'endroit et le moment auxquels elle est visualisée, le tout en temps réel. Le RTB est l'un des modes d'achat programmatique, son mode de commercialisation est aux enchères* »¹⁰. Ces évolutions technologiques accroissent la place des données¹¹ dans le déroulement des

⁴ Cf. par exemple, *The commercial use of consumer data Report on the CMA's call for information*, Competition and Markets Authority, 2015, p.24-25.

⁵ Contrairement aux données structurées, elles ne sont pas organisées selon un modèle prédéfini (ex : formats des données et relations entre les données).

⁶ OCDE, *Data-Driven Innovation, Big Data for Growth and Well-Being*, octobre 2015, p.151.

⁷ Suite à cette invention, Hadoop, une implémentation open-source de MapReduce, a été développée et mise à disposition du public en 2006. Hadoop est aujourd'hui un standard en matière d'exploitation de données massives.

⁸ <https://m.idc.com//pressRelease/prUS40560115>.

⁹ Lexique du Syndicat des Régies Internet accessible à l'adresse suivante : <http://www.sri-france.org>

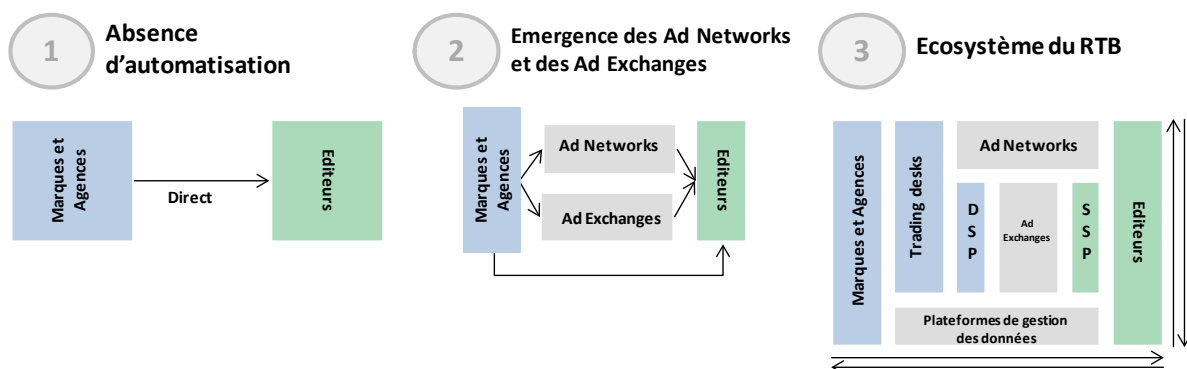
¹⁰ Lexique du Syndicat des Régies Internet.

¹¹ Les principales catégories de données qui sont exploitées sont celles concernant les clients, les données relatives aux offres des annonceurs et les données des médias sur lesquels les annonces sont affichées.

campagnes publicitaires des annonceurs, avec pour objectif principal d'améliorer la performance des campagnes.

6. En France, les modes d'achat programmatique, qui constituent d'ores et déjà le mode d'acquisition principal dans la publicité liée à la recherche, ont par ailleurs représenté 40 % de la publicité *Display*¹² en 2015 (27 % en 2014). Avec 423 millions d'euros, le chiffre d'affaires de l'achat programmatique dans le *Display* est en progression de 61 % en 2015. La croissance globale du marché français de la publicité en ligne a été de 6 %, avec un chiffre d'affaires net de 3,2 milliards d'euros en 2015¹³.
7. Le développement de l'achat programmatique et des systèmes d'enchères en temps réel a également entraîné l'apparition de nouveaux acteurs et de nouveaux métiers liés à l'exploitation de données. Alors qu'initialement l'achat de gré à gré entre les annonceurs et les médias était très développé, de nombreux acteurs intermédiaires sont désormais actifs dans le processus de commercialisation des espaces publicitaires. Les trois schémas ci-après illustrent l'évolution de l'écosystème de la publicité en ligne.

Évolution de l'écosystème de la publicité en ligne



Source : Bulger Partners

8. Dans l'écosystème du Real Time Bidding, la rencontre entre l'offre et la demande d'espaces publicitaires s'effectue ainsi au sein des *Ad Exchanges*, qui sont des places de marché connectées à des *SSP (Supply Side Platforms)*¹⁴ et des *DSP (Demand Side Platforms)*¹⁵. Le rôle des réseaux publicitaires (*Ad Networks*) est quant à lui d'agréger « l'inventaire publicitaire en provenance de différentes sources (supports et *SSP*) dans le but de le commercialiser »¹⁶. Du côté de la demande des annonceurs, les *Trading Desks* et les *DSP* sont les principaux intermédiaires. Les *Trading Desks* sont des agences spécialisées en RTB qui gèrent les campagnes des annonceurs, et peuvent être intégrées à une agence média ou

¹² La publicité Display désigne « la publicité sur le support Internet avec achat d'espaces et insertion d'éléments graphiques ou visuels déclinés sous plusieurs formats (skyscraper, bannières, pavé, habillage de page, interstitiel...). Le mot display a été créé pour faire la distinction entre la publicité classique sur Internet (display) et les campagnes de liens commerciaux (search). » (www.pubdigitale.fr).

¹³ 15^{ème} Observatoire de l'e-pub SRI.

¹⁴ Une *SSP* est une « plateforme technologique d'arbitrage entre différents acheteurs qui permet d'optimiser la commercialisation de l'inventaire publicitaire d'un éditeur. L'éditeur peut y piloter ses ventes impression par impression, en temps réel, et de façon automatisée » (Lexique du Syndicat des Régies Internet).

¹⁵ Une *DSP* est une « plateforme technologique d'optimisation qui permet d'acheter en temps réel, impression par impression et de façon automatisée les inventaires proposés par les différents *SSP* auxquels le *DSP* est connecté » (Lexique du Syndicat des Régies Internet).

¹⁶ Lexique du Syndicat des Régies Internet.

indépendants. Du point de vue des offreurs, les éditeurs sont généralement connectés à des SSP, qui sont elles-mêmes reliées à des *Ad Exchanges*.

9. Cet écosystème se caractérise également par la présence d'entreprises spécialisées dans la fourniture et l'exploitation de données, qui sont utilisées pour améliorer les performances des campagnes. Parmi ces entreprises, il convient de citer les DMP (*Data Management Platforms*), qui sont des plateformes technologiques permettant «*le traitement et la gestion des données anonymes utilisées à des fins de ciblage publicitaire telles que profils sociodémographiques, centres d'intérêt ou intentions d'achat. Ces données peuvent provenir de sources annonceurs/éditeurs (first party) ou de données tierces (third party data)*»¹⁷.
10. Les données tierces peuvent être commercialisées par des entreprises spécialisées qui collectent des données auprès de sources diverses, notamment grâce à leurs propres technologies: des informations publiques (numéros de téléphone, informations publiées sur les réseaux sociaux), mais également des données auprès des pouvoirs publics et d'entreprises tierces (sites web, banques, magasins en ligne, autres courtiers en données).
11. De manière comparable, certains réseaux sociaux ont développé ces dernières années des offres d'accès à leurs données, dont l'utilité dépasse la publicité en ligne. Twitter a ainsi créé une plateforme donnant accès aux flux de tweets et aux données brutes associées. L'entreprise a conclu dans un premier temps des partenariats avec des revendeurs de données (GNIP, Datasift et NTT), qui jouaient un rôle d'intermédiaire avec les entreprises consommatrices de tweets. Après l'acquisition de GNIP en 2014, Twitter a pris la décision de mettre fin à ces partenariats afin d'établir une relation commerciale directe avec les entreprises achetant ces données. En mars 2015, Facebook a lancé une offre d'accès aux données sociales (Facebook Topic Data), dans le cadre d'un partenariat avec Datasift¹⁸.
12. En conclusion, l'importance croissante de l'exploitation de données dans les stratégies commerciales des acteurs a modifié en profondeur les relations entre les offreurs et les demandeurs d'espaces publicitaires, et a entraîné l'apparition et le développement de nombreux acteurs intermédiaires. Ce mouvement devrait être amplifié dans les prochaines années, compte tenu de la croissance continue du volume et de la variété des données, mais aussi de nombreuses innovations technologiques, qu'elles soient le fait de grands groupes internationaux ou d'entreprises nouvelles de taille modeste. Par ailleurs, le développement des outils de blocage des publicités sur internet est susceptible d'inciter les acteurs de la publicité à améliorer le ciblage des annonces et la performance des campagnes. L'enjeu est de préserver le modèle économique de nombreux éditeurs dont les revenus reposent sur la commercialisation de leur inventaire publicitaire.

II. Discussion

13. Aux termes de l'article L. 462-4 du code de commerce, l'Autorité de la concurrence peut prendre l'initiative de donner un avis sur toute question concernant la concurrence. Cet avis est rendu public. Elle peut également recommander au ministre chargé de l'économie ou au ministre chargé du secteur concerné de mettre en œuvre les mesures nécessaires à l'amélioration du fonctionnement concurrentiel des marchés.

¹⁷ Lexique édité par l'IAB France et accessible à l'adresse suivante : <http://www.iabfrance.com/lexique/dmp-data-management-platform>.

¹⁸ <http://lp.datasift.com/rs/068-LBF-492/images/busting-brand-myths-with-facebook-topic-data-by-datasift.pdf>.

14. La présente saisine d'office pour avis a pour objet de permettre à l'Autorité de la concurrence de procéder à une analyse du rôle des données dans les stratégies commerciales des acteurs. Elle ciblera en particulier le secteur de la publicité en ligne.
15. Cette enquête sectorielle fait suite à la publication, le 10 mai 2016, d'une étude conjointe de l'Autorité de la concurrence et du BundesKartellAmt sur les données et leurs enjeux pour l'application du droit de la concurrence. L'étude examine les liens entre les données et le droit de la concurrence et met en évidence les enjeux pour les autorités de concurrence résultant du développement de modèles économiques fondés sur l'exploitation de données massives. Elle se penche plus particulièrement sur le lien entre la détention de données et l'existence ou le renforcement d'un pouvoir de marché et, à cet égard, met en relief l'importance de critères d'analyse tels que l'accessibilité, le volume et la variété des données. Dans le prolongement de cette étude, l'Autorité de la concurrence concentrera son enquête sur les implications concurrentielles de l'exploitation de données dans le secteur de la publicité en ligne. Elle s'intéressera à la définition des marchés pertinents, l'évaluation du pouvoir de marché des acteurs et l'examen de leurs pratiques commerciales.

1. DONNÉES ET MARCHÉS PERTINENTS

16. Les données sont un intrant indispensable, sans lequel la fourniture de biens et de services numériques ne peut être assurée. En tant que telle, l'utilisation de données collectées par une entreprise pour ses propres besoins (autoconsommation) ne peut être assimilée à la fourniture commerciale de produits sur un marché au sens de la pratique décisionnelle des autorités de concurrence. Cependant, si la majorité des données est générée dans le cadre de l'activité des entreprises et utilisée pour leurs propres besoins, des données peuvent être acquises auprès de fournisseurs tiers sous une forme brute ou avec une valeur ajoutée, afin notamment d'améliorer la qualité de service ou de développer de nouveaux services. Dans ce cadre, les données peuvent être considérées comme des produits fournis sur un marché pertinent.
17. L'instruction de l'Autorité de la concurrence portera sur plusieurs questions.
18. Elle analysera le degré de substituabilité entre les différentes formes de publicités en ligne. Son attention se focalisera notamment sur la publicité sur les réseaux sociaux, qui est fondée sur l'exploitation et la monétisation de nombreuses données des individus.
19. En deuxième lieu, l'Autorité de la concurrence formulera des observations sur les marchés pertinents de gros liés aux relations commerciales des acteurs de l'écosystème de la publicité en ligne. Elle s'intéressera également au degré de substituabilité entre les différentes offres commerciales de données qui sont notamment utilisées par les acteurs de la publicité en ligne.
20. Enfin, l'avis portera sur les marchés dits de la protection de la vie privée (adblockers, services de gestion des données personnelles...). Elle recueillera des informations tant sur les services proposés aux consommateurs que sur les relations entre les acteurs de la publicité en ligne et les fournisseurs de services de blocage publicitaire.

2. DONNÉES ET POUVOIR DE MARCHÉ

21. L'Autorité analysera les questions de concentration et d'intégration verticale des acteurs dans le secteur de la publicité en ligne et la place des données dans l'animation de la concurrence. L'instruction cherchera notamment à déterminer si certains acteurs disposent de positions prépondérantes et d'avantages concurrentiels qui seraient liés à leur capacité à exploiter

certaines ensembles de données essentielles à l'animation de la concurrence et qui ne seraient pas reproductibles par un concurrent.

22. Le secteur de la publicité en ligne se caractérise en effet par la présence d'acteurs, susceptibles de disposer de fortes parts de marchés, de revenus en forte croissance, et qui sont actifs sur une part substantielle de la chaîne de production de services publicitaires. En outre, ces entreprises fournissent un ensemble de services connexes, avec ou sans contrepartie financière, qui constituent des sources de production de données susceptibles de renforcer leur position dans le secteur de la publicité en ligne.
23. L'Autorité de la concurrence s'attachera en particulier à appréhender les conséquences, du point de vue de la concurrence, des stratégies de diversification sur d'autres marchés mises en œuvre par des acteurs disposant de positions puissantes et fondées sur leurs capacités d'exploitation des données.
24. Dans le cadre de son analyse, l'Autorité prendra en considération la variété des sources tierces de données disponibles, ainsi que le rôle des algorithmes et des infrastructures informatiques dans l'exploitation des données et le jeu concurrentiel.

3. DONNÉES ET PRATIQUES COMMERCIALES

25. L'Autorité procédera à une analyse des pratiques commerciales ayant un lien avec l'exploitation de données dans le secteur de la publicité en ligne et évaluera dans quelle mesure certaines d'entre elles seraient susceptibles de restreindre le développement d'une concurrence par les mérites.
26. De manière générale, l'Autorité s'intéressera aux conditions qui pourraient être imposées aux annonceurs, aux éditeurs et aux différents acteurs de la chaîne de production des espaces publicitaires.
27. L'Autorité portera ainsi une attention particulière aux problématiques d'accès. Elle examinera notamment les conditions d'accès aux données et aux offres commerciales de données. Elle analysera également les problématiques d'accès des fournisseurs de données aux infrastructures de l'écosystème publicitaire.
28. L'instruction recherchera également l'existence d'éventuelles différences de traitement, qu'elles concernent les intermédiaires techniques, les annonceurs, ou les éditeurs. En effet, des opérateurs verticalement intégrés dans le secteur de la publicité en ligne pourraient être tentés de favoriser leurs propres services et contenus.

29. L’Autorité examinera enfin le niveau de transparence de l’exploitation des données dans le cadre des campagnes publicitaires et ses effets sur la concurrence. Dans son avis sur la publicité en ligne de 2010, l’Autorité a pu envisager qu’un « *refus de garantir un minimum de transparence dans les relations contractuelles* » soit susceptible de constituer une pratique abusive. Dans le cadre de l’enquête, l’instruction se penchera notamment sur le degré de transparence des indicateurs de performance des campagnes, des flux de données et des flux financiers.

Délibéré sur le rapport oral de Mme Clémence Hardy et de M. Mathieu Guennec, rapporteurs, et l’intervention de M. Nicolas Deffieux, rapporteur général adjoint, par M. Bruno Lasserre, président, Mme Élisabeth Flüry-Hérard et M. Emmanuel Combe, vice-présidents.

La secrétaire de séance,
Caroline Chéron

Le président de séance,
Bruno Lasserre