



Droit de la concurrence et données

10 mai 2016

I. INTRODUCTION	3
II. DONNEES	4
1. QU'ENTEND-ON PAR « DONNEES »?.....	4
2. TYPES DE DONNEES	5
a) <i>Types d'informations</i>	5
b) <i>Données structurées et non structurées</i>	6
c) <i>Méthodes de collecte des données</i>	7
3. PROGRES TECHNOLOGIQUE.....	8
4. ROLE DES DONNEES DANS LES ACTIVITES ECONOMIQUES.....	9
a) <i>Améliorations des produits et services</i>	10
b) <i>Exploitation des nouvelles opportunités commerciales</i>	11
c) <i>Des modèles économiques plus ciblés</i>	11
III. ROLE POSSIBLE DES DONNEES DANS L'ANALYSE CONCURRENTIELLE.....	12
1. LES DONNEES COMME SOURCE DE POUVOIR DE MARCHÉ	13
2. DONNEES, TRANSPARENCE DU MARCHÉ ET CONCURRENCE.....	15
3. LES COMPORTEMENTS ANTICONCURRENTIELS LIES AUX DONNEES	18
a) <i>Fusions et acquisitions</i>	18
b) <i>Pratiques d'éviction</i>	20
c) <i>Les données comme vecteur de discrimination tarifaire</i>	23
d) <i>Données, pouvoir de marché et préoccupations en matière de vie privée</i>	25
IV. BIG DATA ET POUVOIR DE MARCHÉ	29
1. LES SOURCES DE POUVOIR DE MARCHÉ DANS LES ACTIVITES EN LIGNE CONSOMMATRICES DE DONNEES	30
a) <i>Marchés multi-faces et effets de réseau</i>	31
b) <i>Multi-hébergement (« multi-homing »)</i>	32
c) <i>Dynamique de marché</i>	34
2. ÉVALUATION DE L'AVANTAGE DERIVE DES DONNEES DANS DES AFFAIRES PASSEES.....	35
a) <i>Sur les marchés « non numériques »</i>	35
b) <i>Sur les marchés « numériques »</i>	38
3. ASPECTS LIES A LA RARETE DES DONNEES	41
a) <i>Sur le caractère non rival des données</i>	42
b) <i>La disponibilité des données numériques</i>	49
4. ASPECTS LIES AU VOLUME ET A LA VARIETE DES DONNEES COLLECTEES	55
a) <i>Présentation</i>	55
b) <i>Discussion</i>	57
V. CONCLUSION	61

Droit de la concurrence et données

I. Introduction

La collecte, le traitement et l'utilisation commerciale des données sont souvent considérés comme des problématiques relevant plus de la protection des données que du droit de la concurrence. Des cas récents témoignent cependant du fait que les autorités de concurrence ont commencé à envisager la possession et à l'utilisation des données du point de vue de la concurrence, même si, en fin de compte, aucune préoccupation n'a été retenue dans les affaires concernées¹.

Les récents développements des marchés numériques ont facilité à l'émergence d'un certain nombre d'entreprises qui réalisent d'importants chiffres d'affaires avec des modèles économiques impliquant la collecte et l'exploitation commerciale de données (souvent personnelles). Certaines de ces entreprises rassemblent une proportion très importante des utilisateurs du service concerné. Le moteur de recherche Google et le réseau social Facebook en sont probablement les exemples les plus marquants. Tandis que de nombreux services fournis par ces entreprises sont présentés comme « gratuits », leur utilisation implique en pratique la collecte d'un certain nombre d'informations personnelles concernant les utilisateurs. Cette situation a soulevé de nouvelles questions, sur le rôle des données dans les relations entre agents économiques, ainsi que sur l'application du droit de la concurrence à ces relations, notamment en ce qui concerne l'évaluation des données en tant que source de pouvoir de marché. Il est important de noter que, bien que ces questions soient souvent examinées en ayant à l'esprit les exemples de Google et de Facebook, elles sont également pertinentes pour de nombreux autres secteurs d'activité. La collecte des données s'opère en effet dans bien d'autres secteurs que ceux des moteurs de recherche, des réseaux sociaux ou de la publicité en ligne et s'étend aujourd'hui à l'énergie, aux télécommunications, à l'assurance,² à la banque

¹ Voir par ex. Commission européenne, « Google/DoubleClick », COMP/M. 4731, en date du 11.03.2008, www.ec.europa.eu/competition/mergers/cases/decisions/m4731_20080311_20682_de.pdf, §§ 359-366; Commission européenne « Facebook/Whatsapp », COMP/M. 7217, en date du 03.10.2014, www.ec.europa.eu/competition/mergers/cases/decisions/m7217_20141003_20310_3962132_EN.pdf, §§ 180-189.

² Voir par ex., UK Financial Conduct Authority, Call for Inputs: Big Data in retail general insurance (2015), <https://www.fca.org.uk/news/call-for-inputs-big-data>.

ou encore au transport³. En outre, le développement des objets connectés devrait conférer aux données un rôle de plus en plus important pour les industries de produits et non plus seulement pour le secteur des services.

La présente étude vise à nourrir le débat en identifiant certains des principaux paramètres et questions à considérer dans l'analyse des interactions entre les données, le pouvoir de marché des entreprises et le droit de la concurrence. Pour cela, il convient tout d'abord de clarifier ce que le mot « données », ou l'expression générique « big data », peut signifier, d'identifier les différents types de données et leurs caractéristiques, de caractériser les différentes méthodes de collecte des données et d'examiner la manière dont les entreprises peuvent les utiliser (**partie II**). Les scénarios anticoncurrentiels qui peuvent être associés à la collecte et à l'exploitation des données sur les marchés numériques sont ensuite présentés dans la **partie III**. Enfin, c'est à la lumière de ces développements que la **partie IV** discute de certains des paramètres devant être considérés dans l'évaluation de la pertinence et de la crédibilité de tels scénarios anticoncurrentiels.

II. Données

1. Qu'entend-on par « données » ?

Il n'existe pas de définition unique du mot « données ». Dans une acception étroite, ce terme est souvent utilisé pour nommer les résultats des expériences ou des mesures scientifiques. Mais dans un sens plus large, il est employé pour faire référence à une information (quelconque) ou sa représentation, souvent en association avec son stockage sur un ordinateur⁴.

En ce qui concerne plus spécifiquement la concurrence et l'économie numérique, le terme de « **big data** » est souvent employé sans, lui non plus, avoir de définition commune⁵. Les caractéristiques du « big data », qui sont généralement mentionnées ont trait à l'importance, en termes de volume et de variété, des données, à leur mode

³ Voir par ex., Comité du débat sur l'ouverture des données liées à l'offre de transport, présidé par Francis Jutand, rapport (2015), <http://www.ladocumentationfrancaise.fr/rapports-publics/154000182/> – disponible uniquement en français.

⁴ <http://www.merriam-webster.com/dictionary/data>;
<http://www.oxforddictionaries.com/definition/learner/data>.

⁵ Voir Hu, Han et al., Toward Scalable Systems for Big Data Analytics: A Technology Tutorial, IEEE Access, Vol. 2 (2014) ; <http://ieeexplore.ieee.org/xpl/articleDetails.jsp?reload=true&arnumber=6842585>, p. 652; Gil Press, 12 Big Data Definitions: What's Yours?, Forbes, <http://www.forbes.com/sites/gilpress/2014/09/03/12-big-data-definitions-whats-yours>; The Big Data Conundrum: How to Define It?, MIT Technology Review, <http://www.technologyreview.com/view/519851/the-big-data-conundrum-how-to-define-it/>.

de collecte (à grande vitesse à partir de sources multiples), ainsi qu'à la puissance informatique et algorithmique requise pour leur traitement et leur analyse⁶. Ainsi, le « big data » se caractérise généralement par les trois « V », pour *Vitesse*, *Variété* et *Volume*, auxquels un quatrième « V » peut être ajouté, celui de la *Valeur* qui peut être extraite des données⁷. Au final, si le terme « big data » est utilisé de différentes manières et se fonde sur des caractéristiques aux contours flous, il tente néanmoins de rendre compte des récents développements qu'ont connus les marchés et les technologies numériques ces deux dernières décennies.

2. Types de données

Il existe plusieurs approches permettant de classer les données.

a) Types d'informations

Tout d'abord, les données peuvent être classées en fonction du **type d'informations** qu'elles apportent, qui peuvent varier, pour une entreprise donnée, en termes d'utilité ou de disponibilité. Les données peuvent fournir des informations sur des individus, des entités économiques ou des objets ; elles peuvent par exemple révéler des informations sur le comportement, les préférences ou la localisation géographiques des personnes, sur le volume d'affaires réalisé par une entreprise avec certaines transactions, ou enfin sur la localisation et la vitesse d'une voiture à un moment donné.

Jusqu'à présent, le débat qui concerne la relation entre les données et le droit de la concurrence se concentre cependant principalement sur les données dites personnelles⁸. Ces données personnelles font l'objet de règles spécifiques, qui contraignent leur collecte, leur traitement et leur utilisation aux fins de protéger la vie privée des personnes auxquelles elles se rattachent. Les données personnelles sont définies à l'Article 2 (a) de la Directive 95/46/CE (Directive relative à la protection des

⁶ European Data Protection Supervisor, https://secure.edps.europa.eu/EDPSWEB/edps/Consultation/big_data.

⁷ Voir Hu, Han et al., Toward Scalable Systems for Big Data Analytics: A Technology Tutorial, IEEE Access, Vol. 2 (2014), <http://ieeexplore.ieee.org/xpl/articleDetails.jsp?reload=true&arnumber=6842585>, p. 652 (654); German Monopolies Commission (Monopolkommission), Special Report No. 68: Competition policy: The challenge of digital markets (2015), http://www.monopolkommission.de/images/PDF/SG/s68_fulltext_eng.pdf, § 67.

⁸ Selon les agences européennes de protection des données personnelles, le champ des données à caractère personnel s'étend aux données qui peuvent être attribuées à une adresse IP et à des cookies, même si le nom de l'utilisateur n'est pas identifié. Voir Article 29 Data Protection Working Party, Opinion 1/2008 on data protection issues related to search engines, en date du 4 April 2008, http://ec.europa.eu/justice/policies/privacy/workinggroup/wpdocs/index_en.htm, pp.6-8, et Opinion 4/2007 on the concept of personal data, en date du 20th June 2007, Voir lien ci-dessus, pp.16-17.

données) comme « *toute information concernant une personne physique identifiée ou identifiable* ». Les données à caractère personnel ne peuvent être collectées et traitées que de manière limitée. L'Article 6 (1) (b) de la Directive relative à la protection des données dispose ainsi que les données à caractère personnel doivent être « *collectées pour des finalités déterminées, explicites et légitimes, et ne pas être traitées ultérieurement de manière incompatible avec ces finalités* ». Par ailleurs, le traitement des données à caractère personnel requiert une base juridique spécifique, qui peut notamment impliquer le « *consentement indubitable* » de la personne concernée (Art. 7 (a) de la Directive relative à la protection des données), entendu comme la manifestation de sa volonté, libre, spécifique et informée par laquelle elle accepte que des données à caractère personnel la concernant fassent l'objet d'un traitement (Art. 2 (h) Directive relative à la protection des données).

b) Données structurées et non structurées

Ensuite, les données peuvent être classées en fonction de leur caractère « **structuré** » ou non. Cette distinction a en effet des conséquences sur les possibilités de valorisation des données et sur les méthodes pour y parvenir.

Les données structurées le sont au travers d'un certain nombre de champs, du type d'information que ces champs contiennent et de la manière dont ces champs sont liés les uns aux autres. Une base de données d'adresses de consommateurs contenant les informations relatives au prénom, au nom, à l'adresse, à l'âge, aux numéros de téléphone, etc. du consommateur est un exemple de données structurées. Les données structurées sont ainsi plus faciles à traiter et à utiliser à des fins commerciales que les données non structurées, au regard tout du moins des méthodes traditionnelles d'exploitation de données. Les données non structurées ne se conforment en effet pas à un modèle déterminé et, de ce fait, nécessitent en général d'être traitées par des algorithmes plus récents pour qu'une valeur commerciale puisse en être extraite⁹. Les données peuvent également être définies comme semi-structurées lorsque leurs caractéristiques relèvent des deux catégories précédentes. Dans ce cas, elles ne relèvent pas d'un modèle prédéfini, mais certains

⁹ Voir Jim Harris, Bridging the Divide between Unstructured and Structured Data, <https://datascience.berkeley.edu/structured-unstructured-data/>; Stephen Pritchard, How to manage unstructured data for business benefit, ComputerWeekly, <http://www.computerweekly.com/feature/How-to-manage-unstructured-data-for-business-benefit> ; Bill Schoonmaker, Unstructured Data Can Create Chaos, Forbes, www.forbes.com/sites/emc/2013/07/11/unstructured-data-can-create-chaos/ .

éléments ou champs qu'elles incluent peuvent être identifiés de manière automatisée.

c) Méthodes de collecte des données

Enfin, les données peuvent être classées selon la méthode par laquelle elles ont été collectées. Les données peuvent en effet être collectées de différentes manières susceptibles d'influencer leur valeur économique, notamment en lien avec la question de leur éventuelle rareté.

Les données sont souvent fournies activement et de manière **volontaire** par les individus. Une boutique en ligne demandera par exemple au consommateur de saisir ses adresses postales et électroniques ainsi que ses coordonnées de paiement afin de pouvoir traiter la transaction. Les réseaux sociaux comme les services de communication du même type accèdent quant à eux à toutes sortes de données communiquées par leurs utilisateurs (principalement des données personnelles). Ces données peuvent inclure des informations personnelles comme le nom, l'adresse ou le niveau de diplôme, ainsi que le contenu des messages, des photos, des vidéos, des commentaires sur l'actualité, des habitudes de consommation, etc. Les moteurs de recherche, généraux ou spécialisés, exploitent les termes de recherche saisis par leurs utilisateurs qui révèlent des informations sur leurs centres d'intérêt ; une plateforme vidéo a besoin que ses utilisateurs partagent de nouveaux contenus tandis qu'une plateforme de rencontre ne peut rapprocher les personnes que si elles fournissent suffisamment d'informations et de détails sur leur vie privée.

Les données peuvent également être rassemblées tout simplement en exploitant les sources (librement) accessibles sur internet ou en **observant** le comportement des internautes, même à leur insu. Dans le premier cas, on peut citer les moteurs de recherche (généraux) qui exploitent l'analyse systématique qu'ils font de chaque page disponible sur le web, technique également connue sous la désignation de « crawling » (indexation). Dans le second cas, également très répandu, figure le grand nombre d'entreprises qui « suivent » les internautes au cours de leur navigation en ligne d'une page à une autre, et même parfois à l'intérieur d'une seule et même page¹⁰, à l'aide de différentes techniques. Ces techniques se sont

¹⁰ Qu'il soit techniquement possible de surveiller la partie d'une page web qu'un utilisateur voit effectivement a suscité des demandes de la part des annonceurs d'un minimum de « visibilité » des publicités *display* pour leur prise en compte aux fins de mesurer l'audience (les pages web sont en général plus grandes que les écrans sur lesquels elles sont visionnées) Voir Invisible ads,

développées afin de pouvoir fonctionner sur différents appareils et non exclusivement sur Internet mais également sur des applications mobiles, etc.¹¹. Certains de ces outils peuvent être contournés par les utilisateurs de manière assez simple¹², mais pour d'autres cela peut être très difficile, voire impossible. Les données ainsi générées peuvent ensuite être facilement associées (techniquement) aux données volontairement renseignées par un individu, afin de catégoriser de manière très fine les utilisateurs. La législation sur la protection des données personnelles peut cependant limiter les possibilités dans ce domaine.

Enfin, de nouvelles informations peuvent être générées par **déduction** à partir de données existantes. Une boutique de mode en ligne peut par exemple analyser les produits qu'un visiteur a regardés pour en déduire, avec une certaine marge d'erreur, que le visiteur en question est un homme ou une femme. Une entreprise proposant différents services en ligne peut combiner les données des utilisateurs de ses services afin d'obtenir de nouvelles informations les concernant.

3. Progrès technologique

Les développements récents sur les marchés numériques permettent le traitement d'une quantité croissante d'informations tout en réduisant le temps nécessaire à ce traitement. Un certain nombre de traitements complexes et longs à mettre en œuvre n'ont ainsi été rendus possibles que par certaines évolutions techniques.

L'un des principaux facteurs de ce développement a été l'augmentation continue de la puissance des ordinateurs au cours des dernières décennies. Selon une règle empirique connue sous le nom de « loi de Moore », le nombre de transistors dans un circuit intégré double environ tous les deux ans¹³. Le nombre croissant de transistors sur un microprocesseur, ainsi que la diminution de leur coût, ont alors permis

phantom readers, The Economist, 26th March 2016, <http://www.economist.com/news/business/21695388-worries-about-fraud-and-fragmentation-may-prompt-shake-out-crowded-online-ad> .

¹¹ Voir Dan Goodin, Beware of ads that use inaudible sound to link your phone, TV, tablet, and PC, ars technica, <http://arstechnica.com/tech-policy/2015/11/beware-of-ads-that-use-inaudible-sound-to-link-your-phone-tv-tablet-and-pc/> .

¹² Même si un grand nombre d'utilisateurs semblent ne pas le savoir ou ne pas y attacher d'importance.

¹³ Il y a actuellement une discussion sur l'avenir de ce développement, Voir par ex., Peter Bright, Moore's law really is dead this time, ars technica, <http://arstechnica.com/information-technology/2016/02/moores-law-really-is-dead-this-time/> ; M. Mitchell Waldrop, The chips are down for Moore's law, Nature, <http://www.nature.com/news/the-chips-are-down-for-moores-law-1.19338> ; After Moore's law – The future of computing, The Economist, 12.03.2016, <http://www.economist.com/news/leaders/21694528-era-predictable-improvement-computer-hardware-ending-what-comes-next-future> .

d'obtenir des microprocesseurs plus rapides et une capacité plus importante de mémoire vive. Cet accroissement des capacités et de la vitesse s'observe également dans les technologies utilisées pour stocker les données de manière permanente (telles que les disques durs et les mémoires flash)¹⁴.

Un autre facteur a été l'augmentation massive de la vitesse et de la capacité des réseaux de transport de données d'un point à un autre. La comparaison de ce qu'un foyer moyen avait à sa disposition à la fin des années 1980¹⁵ et de ce dont il peut disposer aujourd'hui¹⁶ en est l'illustration. Cette constatation s'applique également aux réseaux utilisés par les entreprises ainsi qu'à toute l'infrastructure de télécommunications.

Un troisième facteur qui doit être pris en compte est le développement de nouvelles méthodes permettant d'extraire des informations valorisables à partir de volumes extrêmement importants de données (souvent non structurées). Sans ces développements, un service comme le moteur de recherche Google n'aurait pas été possible.

4. Rôle des données dans les activités économiques

Le fait que les entreprises recourent aux données dans leur activité n'est pas un phénomène nouveau. Même dans la « vieille économie », les données des clients étaient une source essentielle d'informations pour les entreprises, notamment pour s'adresser aux clients potentiels au travers de la publicité, pour prédire leurs préférences d'achat ou leurs aptitudes à dépenser et pour analyser les performances de leurs salariés. Le marketing est fondé sur la prospection de marchés, ce qui inclut la collecte, le traitement, et l'analyse systématiques de données. Chaque entreprise souhaite avoir une connaissance aussi complète que possible des préférences de ses clients afin d'améliorer ses produits, d'offrir des services personnalisés et de cibler ses annonces publicitaires.

¹⁴ Voir par ex., Timeline: 50 Years of Hard Drives, PCWorld, <http://www.pcworld.com/article/127105/article.html>; et la liste des ressources sur <http://www.storagenewsletter.com/rubriques/miscellaneous/history-of-storage-industry/>.

¹⁵ Le service qu'un consommateur pouvait en général obtenir à la fin des années 80 et au début des années 90 était un modem livrant quelques milliers de bits par seconde, Voir Tamsin Oxford, Getting connected: a history of modems, Techradar, <http://www.techradar.com/news/internet/getting-connected-a-history-of-modems-657479/1> ; Benj Edwards, 60 Years of Hooking Up, PCWorld, http://www.pcworld.com/article/218274/modems_through_the_years.html#slide1.

¹⁶ La plupart des foyers en Allemagne obtiennent par exemple quelques millions de bits par seconde, et dans certaines zones cent ou deux cent millions de bits par seconde.

Les évolutions technologiques de l'économie numérique ont cependant révolutionné les possibilités de collecte, de traitement et d'utilisation commerciale des données dans presque tous les secteurs d'activité. C'est la raison pour laquelle une attention particulière est maintenant portée aux avantages économiques qui peuvent être tirés de l'utilisation des données par une entreprise. Quand bien même la collecte et l'utilisation de volumes importants de données sont susceptibles dans certains cas de soulever des préoccupations de concurrence (voir Parties III et IV), il s'agit également d'un mécanisme par lequel les produits et services peuvent être améliorés et l'efficacité économique accrue.

a) Améliorations des produits et services

Les données peuvent contribuer à améliorer les produits ou services d'une entreprise. Cela peut être la conséquence d'un effet d'apprentissage, comme dans le cas des moteurs de recherche sur Internet. On peut en effet raisonnablement supposer que l'augmentation du nombre de recherches, couplée à la possibilité d'observer les résultats sur lesquels chaque utilisateur clique, permettent au moteur de recherche d'améliorer et d'affiner ses résultats ainsi que le fonctionnement de son algorithme. L'accroissement de la qualité qui en résulte peut ensuite inciter davantage de personnes à utiliser ce moteur de recherche. De la même manière, de nombreux logiciels installés sur les ordinateurs personnels ou les smartphones collectent des informations détaillées concernant l'utilisation de ces produits. Des exemples frappants sont donnés par les navigateurs web¹⁷ et les systèmes d'exploitation¹⁸. Par ailleurs, de nombreux sites Internet, sinon la plupart, rassemblent également des informations détaillées sur la navigation de leurs utilisateurs et utilisent ces informations afin d'identifier les parties du site qui ont été consultées de manière plus intensive ou de réduire les problèmes techniques. Ces informations peuvent par exemple être utilisées pour développer les parties du site Internet qui sont le plus souvent consultées ou accélérer les fonctions les plus utilisées d'un logiciel afin de l'améliorer.

¹⁷ Par ex., Firefox, si l'utilisateur a adhéré : <https://support.mozilla.org/en-US/kb/share-telemetry-data-mozilla-help-improve-firefox> .

¹⁸ Par ex., « Windows 10 » de Microsoft, dans lequel le retrait complet n'est pas possible pour les utilisateurs privés. Voir Woody Leonhard, Microsoft walks a thin line between Windows 10 telemetry and snooping, InfoWorld, <http://www.infoworld.com/article/3020152/microsoft-windows/microsoft-walks-a-thin-line-between-windows-10-telemetry-and-snooping.html> ; Nick Heath, Windows 10 now lets you turn off tracking - but only if you're a business, TechRepublic, <http://www.techrepublic.com/article/windows-10-now-lets-you-turn-off-tracking-but-only-if-youre-a-business/> .

Au-delà des exemples décrits ci-dessus, il est clair que l'accroissement des volumes de données peut être particulièrement déterminant pour les services où les données constituent plus ou moins le produit lui-même ou sont, au minimum, étroitement liées au produit. C'est le cas des « plateformes de matching » comme les sites de rencontres en ligne : plus les utilisateurs en disent sur leur profil personnel, plus le service devient attrayant pour ceux qui cherchent à trouver le partenaire idéal. YouTube en tant que plateforme vidéo est également d'autant plus attractive qu'elle propose davantage de vidéos. Cependant, dans les deux cas, l'amélioration du service dépend en partie de la capacité de la plateforme à permettre une navigation aisée au sein de cette mine de données et à préserver l'expérience de l'utilisateur, tout en augmentant l'étendue et la profondeur de son offre.

b) Exploitation des nouvelles opportunités commerciales

L'accès aux données peut également permettre aux entreprises d'exploiter de nouvelles opportunités commerciales. Les entreprises peuvent ainsi réutiliser les données collectées à l'occasion de la fourniture d'un service en vue de proposer de nouveaux services basés sur ces données. C'est ainsi, par exemple, que les données de mobilité générées par les opérateurs de réseau mobile et les terminaux sont utilisées par des prestataires de service de navigation cherchant à améliorer l'information sur les encombrements routiers et à orienter leurs utilisateurs de manière à les éviter¹⁹.

c) Des modèles économiques plus ciblés

Les données peuvent également être utilisées pour mieux cibler les clients potentiels et leur fournir des messages publicitaires, des services ou des produits personnalisés. Les entreprises peuvent ainsi réduire leurs coûts publicitaires en restreignant leur campagne au public qu'elles ciblent réellement. Les entreprises peuvent également fixer des prix individualisés en fonction des aptitudes de dépenses estimées des consommateurs ainsi que sur leur sensibilité au prix. Quand bien même un tel mode de personnalisation des prix n'est pas encore très développé²⁰, il pourrait constituer un enjeu à l'avenir.

¹⁹ Voir : Gianna-Carina Grün, Der beste Staumelder ist das eigene Handy, Die Zeit, <http://www.zeit.de/digital/mobil/2012-06/staudaten-handy> – disponible exclusivement en allemand.

²⁰ Voir le rapport conjoint de la CNIL et de la DGCCRF, <http://www.cnil.fr/linstitution/actualite/article/article/ip-tracking-conclusions-de-lenquete-conjointe-menee-par-la-cnil-et-la-dgccrf/> ainsi que le rapport de l'OFT, Personalised Pricing: Increasing Transparency to Improve Trust,

La publicité en ligne basée sur le « ciblage comportemental » peut servir d'exemple à un modèle d'affaires qui n'a été rendu possible, tel qu'il existe actuellement, que grâce aux développements techniques qui viennent d'être décrits. Le « ciblage comportemental » caractérise le fait de délivrer des publicités en ligne à des utilisateurs spécifiques en fonction de leur profil constitué à partir de l'observation de leurs habitudes de navigation²¹. Ce modèle d'affaires n'existerait pas sans les développements mentionnés pour trois raisons. Tout d'abord, sans la pénétration du numérique et des technologies de la communication dans l'essentiel, sinon la totalité, des domaines de la vie, il ne serait pas possible de suivre le comportement d'un aussi grand nombre de personnes de manière aussi détaillée. Ce n'est ainsi que depuis qu'une grande partie de la population a commencé à lire les actualités en ligne, à voir des films et des émissions en ligne, à acheter en ligne, à écouter de la musique en flux (streaming) ou des stations de radio en ligne, à commenter la politique ou la dernière tenue des célébrités en ligne, à se présenter sur les réseaux sociaux et à avoir sur soi un appareil qui permette d'enregistrer sa position géographique, qu'il a été possible d'enregistrer leurs actes de manière si précise que des informations détaillées et personnalisées les concernant peuvent en être tirées. Deuxièmement, afin de mettre cela en œuvre à très grande échelle, une grande puissance informatique et d'importantes capacités de stockage de données sont nécessaires. Enfin, l'omniprésence des canaux de communications numériques et leur rapidité sont nécessaires à la diffusion de ces messages individualisés.

III. Rôle possible des données dans l'analyse concurrentielle

L'influence des données comme facteur de compétitivité des entreprises attire d'autant plus l'attention que les entreprises collectent de plus en plus d'informations sur leurs utilisateurs et leurs préférences. La jurisprudence et la littérature identifient

http://webarchive.nationalarchives.gov.uk/20140402142426/http://www.ofcom.gov.uk/shared_ofcom/markets-work/personalised-pricing/ofcom1489.pdf .

²¹ Voir [smm Sales & Marketing Management: Behavioral Targeting, https://www.salesandmarketing.com/article/behavioral-targeting](https://www.salesandmarketing.com/article/behavioral-targeting); What Is Behavioral Targeting?, CBS News, <http://www.cbsnews.com/news/what-is-behavioral-targeting/>.

les problématiques suivantes. Tout d'abord, la collecte et l'exploitation des données peuvent créer des barrières à l'entrée et être une source de pouvoir de marché (1). Ensuite, elles peuvent également accroître la transparence du marché et donc influencer son fonctionnement (2). Plusieurs types de pratiques liées aux données d'une entreprise peuvent enfin soulever des préoccupations de concurrence (3).

1. Les données comme source de pouvoir de marché

Lorsque l'accès à un large volume ou à une importante variété de données est un facteur de compétitivité sur le marché (ce qui est spécifique au marché en question), leur collecte peut constituer une barrière à l'entrée si de nouveaux entrants ne sont pas en mesure de collecter ou d'acheter le même type de données, en termes de volume et/ou de variété, que les entreprises déjà en place.

Comme évoqué dans la partie II, il existe différentes manières de rassembler les données. Elles peuvent être fournies volontairement par les utilisateurs ou les prospects d'une entreprise. Elles peuvent également être observées par cette entreprise lorsqu'elle interagit avec un utilisateur (ou un prospect) ou encore être déduites de cette interaction, ce qui est par exemple le cas lorsqu'un distributeur en ligne observe les articles sur lesquels un client a cliqué sans toutefois les avoir achetés. Dans tous ces cas, l'entreprise contrôle elle-même la collecte des données car elle est impliquée dans la relation avec l'utilisateur (ou le prospect) (« données de première main »). Disposant d'un nombre inférieur d'utilisateurs, les entreprises de plus petite taille, ou les nouveaux entrants, peuvent en général collecter moins de données de « première main » que leurs concurrentes plus grandes et mieux établies.

Cependant, les entreprises peuvent également utiliser des « données de tierce partie » en plus des données de « première main », c'est-à-dire des données collectées par une autre entité qu'elles-mêmes, si toutefois celles-ci sont accessibles. Il peut en effet y avoir des limites à la disponibilité des données de tierce partie, lorsque par exemple l'entreprise qui les collecte n'est pas disposée à les partager ou à les vendre à des concurrents. En ce qui concerne leurs coûts d'accès, les coûts fixes associés aux données de tierce partie peuvent être inférieurs à ceux des données de première main, mais les coûts variables peuvent s'avérer supérieurs. Surtout, le recours à des données de tierce partie permet, au moins en principe, de

créer des bases de données qui ne limitent pas aux clients (prospects) de l'entreprise.

Mais si chaque entreprise peut donc en théorie acquérir des « données de tierce partie », et ainsi rattraper l'avantage des entreprises établies en matière d'accès aux données, en pratique cela pourrait s'avérer impossible au regard de la quantité et de la qualité des données déjà détenues par l'entreprise en place. Dans certains secteurs, les principales entreprises peuvent disposer d'une base de clientèle et d'informations tellement large que la question de savoir si une entreprise tierce est en mesure de reproduire le même volume et la même variété de données se pose. Cela peut être notamment (mais pas exclusivement) le cas de services en ligne tels que les moteurs de recherche ou les réseaux sociaux où des services « gratuits » particulièrement attractifs sont proposés à une large base d'utilisateurs qui, en retour, génèrent un volume important de données susceptibles ne pas être accessibles aux concurrents.

L'action engagée par le DoJ en 2014 contre la fusion de *Bazaarvoice* et de son principal rival *Power-Reviews* a par exemple établi que les données peuvent constituer une barrière à l'entrée sur le marché des « plateformes de notation et d'avis d'utilisateurs ». La fusion horizontale entre ces deux entreprises aurait réduit la concurrence entre les « plateformes de notation et d'avis d'utilisateurs » aux États-Unis en créant un quasi-monopole sur ce marché, amplifié par des barrières significatives à l'entrée, y compris résultant des effets de réseau liés à l'association de contenus, des coûts de transfert, de modération, d'analyses et de réputation²². Un tel accroissement des barrières à l'entrée pourrait protéger les entreprises en place de leurs rivaux plus modestes et d'entrants potentiels, et leur permettre ainsi d'augmenter leurs prix. En soi, la perspective de barrières plus importantes à l'entrée, qui doit toujours être évaluée au cas par cas comme l'explique la partie IV, n'a de conséquences significatives sur la concurrence que si le niveau de concentration du marché est relativement élevé, ou si les caractéristiques du marché sont favorables à une collusion tacite. Dans ce cas, en dépit du fait que les consommateurs et l'efficacité économique bénéficient des gains de productivité associés au développement de la collecte des données et des usages possibles,

²² DOJ, Antitrust Division, Competitive Impact Statement, 13-cv-00133 WHO, en date du 08.05.2014, <http://www.justice.gov/atr/case-document/file/488826/download>, p. 5.

l'accroissement des barrières à l'entrée et la réduction de concurrence qui en résulte pourraient s'avérer dommageables pour l'économie.

Il apparaît que les secteurs de l'économie où la collecte et l'utilisation des données sont souvent considérées comme particulièrement importantes, à l'instar par exemple des moteurs de recherches ou des réseaux sociaux, s'avèrent souvent concentrés, quelques opérateurs concentrant une part importante des utilisateurs. En outre, l'existence alléguée d'économies d'échelle et d'effets de réseau peuvent également compter comme facteurs de limitation de l'intensité concurrentielle (voir partie IV). Le développement de la collecte et de l'utilisation des données sur ces marchés peut ainsi renforcer le pouvoir de marché des entreprises en place. La marginalisation des concurrents plus modestes ayant un accès différencié aux données pourrait en effet aller dans ce sens puisque l'accès à des volumes plus importants de données peut permettre d'offrir de meilleurs services, qui, à leur tour, permettent d'attirer davantage d'utilisateurs et donc de collecter davantage de données (« effet boule de neige »). A contrario, les petites entreprises n'ont pas autant d'utilisateurs et récupèrent en conséquence moins de données. A mesure que l'écart en part de marché se creuse, l'écart des capacités de collecte de données pourrait donc également se creuser, aggravant l'écart de qualité des services proposés aux utilisateurs. Enfin, les revenus supérieurs engrangés par les grandes entreprises permettent de financer des investissements plus importants (comme de nouveaux algorithmes, de nouvelles fonctionnalités, l'entrée sur des marchés adjacents, etc.), attirant ainsi encore plus d'utilisateurs et donc de données. Cette tendance pourrait donc être préjudiciable à la concurrence en faisant converger les marchés liés aux données vers une situation de monopole²³.

2. Données, transparence du marché et concurrence

La collecte et l'utilisation croissantes des données numériques sont souvent associées à un accroissement de la transparence des marchés en ligne. D'un point de vue économique, cela a des effets ambigus sur le fonctionnement de ces marchés.

Les consommateurs peuvent, d'une part, bénéficier d'une meilleure transparence du marché si cela leur permet de comparer plus facilement les prix ou les

²³ Dans certains cas, le développement des données peut également réduire les obstacles à l'entrée, par exemple lorsque les données acquises sur un marché donné peuvent être utilisées pour identifier et satisfaire les besoins des consommateurs sur un autre marché.

caractéristiques de biens ou services. Les comparateurs de prix ou les sites comme Tripadvisor permettent par exemple aux consommateurs de faire des choix mieux informés, ce qui se traduit par une concurrence plus vive tant en termes de prix que de qualité. Les places de marché sont un autre exemple illustrant les avantages que la transparence du marché peut apporter. Amazon ou E-Bay hébergent ainsi de nombreuses boutiques en ligne, y compris les plus petites qui, sans cette plateforme, auraient pu avoir des difficultés à entrer sur le marché. En outre, ces plateformes permettent de comparer les prix et les conditions offertes par leurs commerçants hébergés, contribuant ainsi à rendre le marché plus transparent. Dans certains cas, cet accroissement de la transparence peut également faciliter l'entrée de nouveaux concurrents car ils ont plus d'informations sur les besoins des consommateurs et sur les conditions d'accès au marché.

D'autre part, l'accroissement du niveau d'information du fait du développement de la collecte des données, notamment sur les prix pratiqués par les concurrents, peut également être utilisé par les entreprises pour limiter la concurrence. La transparence du marché peut en effet favoriser la stabilité d'une collusion (tacite ou explicite). En permettant de détecter plus facilement la déviation par rapport à un accord collusif, la transparence du marché limite le profit escompté et par conséquent diminue l'incitation d'une partie à remettre en cause l'équilibre collusif. Dans une certaine mesure, la disponibilité croissante des données de prix sur Internet, et le fait que ces données sont affichées en temps réel, pourraient donner aux marchés en ligne un niveau de transparence jamais atteint jusqu'alors²⁴.

La collecte des données peut également faciliter la collusion lorsque ces données sont utilisées pour fixer les prix au travers de l'utilisation d'algorithmes. Même si la transparence du marché, en tant que facteur de facilitation de la collusion, est l'objet de débats depuis maintenant plusieurs décennies, les développements techniques en particulier dans le domaine des algorithmes informatiques rendent ces débats encore plus pertinents. C'est ainsi, par exemple, qu'en traitant toutes les informations disponibles et en contrôlant, analysant ou en anticipant les réponses de la concurrence aux prix actuels et futurs, des concurrents peuvent plus facilement mettre en œuvre un équilibre tarifaire supra-concurrentiel durable²⁵. En outre, les

²⁴ Bien que l'utilisation de système de tarification personnalisée puisse contrer cet effet dans une certaine mesure.

²⁵ Des exemples réels de ces scénarios incluent *United States v. Airline Tariff Publ'g Co.*, 836 F. Supp. 9, 12 (D.D.C. 1993) ainsi que l'affaire *Topkins*. La décision a dans cette affaire

algorithmes qui traitent les données peuvent également être utilisés pour mettre en œuvre une entente, détecter les déviations et permettre plus généralement aux prix collusifs de réagir de manière plus précise aux évolutions des conditions exogènes du marché.

Enfin, certains analystes²⁶ ajoutent également à ces risques deux préoccupations plus sophistiquées. Tout d'abord, même en l'absence de coordination horizontale explicite, l'emploi d'algorithmes tarifaires similaires, ce qui peut par exemple être le cas si ces algorithmes sont fournis par la même société, pourrait porter atteinte à la concurrence en réduisant l'incertitude et les biais comportementaux qui sont favorables à la concurrence par les prix. Ensuite, les algorithmes peuvent également limiter la concurrence en intégrant dans leurs mécanismes de fixation des prix les réactions des concurrents pouvant être déduites des données collectées au cours des expériences passées de variations des prix. Ainsi, même des algorithmes d'origines différentes pourraient servir, par exemple, à suivre les augmentations tarifaires des concurrents, à sanctionner des écarts, etc. Alternativement, une collusion tacite pourrait également découler d'un processus automatisé d'acquisition des connaissances. Pour autant, il peut s'avérer difficile d'engager des poursuites à l'encontre de telles pratiques. Tout d'abord, la transparence du marché est généralement considérée comme bénéfique aux consommateurs, au moins en théorie, lorsqu'ils accèdent aux mêmes informations que les entreprises. Ensuite, il est possible qu'aucune coordination ne soit nécessaire pour obtenir les mêmes résultats supra-concurrentiels auxquels aboutirait une collusion anticoncurrentielle.

(http://www.justice.gov/sites/default/files/opa/press-releases/attachments/2015/04/06/topkins_information.pdf et <http://www.justice.gov/opa/pr/former-e-commerce-executive-charged-price-fixing-antitrust-divisions-first-online-marketplace>) indiqué que « afin de mettre en œuvre cet accord, TOPKINS et ses co-conspirateurs ont adopté des algorithmes tarifaires spécifiques des affiches convenues dans l'objectif de coordonner les évolutions de leurs prix respectifs. Pour servir la conspiration, TOPKINS a écrit le code informatique du logiciel basé sur un algorithme de la Société A pour qu'il fixe les prix des affiches convenues conformément à cette entente. Aux fins d'aboutir à des accords tarifaires, de mettre en œuvre le respect des accords obtenus, et de contrôler l'efficacité des algorithmes tarifaires, TOPKINS et ses co-conspirateurs ont échangé, contrôlé et discuté des informations sur les prix et les ventes des affiches convenues. Conformément aux accords conclus, TOPKINS et ses co-conspirateurs ont vendu, distribué, et accepté le paiement des affiches convenues à des prix collusifs non concurrentiels sur Amazon. »

²⁶ Voir Maurice E. Stucke and Ariel Ezrachi, Artificial intelligence and collusion: when computers inhibit competition, University of Tennessee, Legal Studies Research Paper No. 267 (2015), http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2591874.

3. Les comportements anticoncurrentiels liés aux données

Lorsque les données confèrent des avantages concurrentiels significatifs à leurs propriétaires, les entreprises doivent acquérir davantage de données et/ou mieux les analyser et les exploiter pour rester compétitives et/ou obtenir un avantage compétitif par rapport à leurs rivaux sur le marché. Les pratiques mises en œuvre pour acquérir ces données peuvent alors être examinées au travers du prisme du droit de la concurrence. Parallèlement, les entreprises peuvent utiliser leur pouvoir de marché basé sur les données pour obtenir un avantage compétitif sur des marchés adjacents. Des préoccupations de concurrence sont également souvent soulevées lorsque les données permettent de discriminer en prix. Enfin, les interactions possibles entre le droit de la concurrence et les règles de protection de la vie privée sont examinées ci-après.

a) Fusions et acquisitions

Pour obtenir un meilleur accès aux données, une des premières stratégies envisageable pour une entreprise est d'acquérir d'autres entreprises possédant de vastes bases de données ou de fusionner avec elles. L'OCDE (2015)²⁷ indique à cet égard que, dans les secteurs liés aux données, « *le nombre des fusions et acquisitions (M&A) a rapidement augmenté de 55 opérations en 2008 à près de 164 en 2012* » (p. 94).

Sur de nombreux marchés, la fusion entre une entreprise bien établie et un nouvel entrant innovant n'a qu'un faible impact sur la structure du marché en raison des faibles parts de marché de l'entrant ou de l'absence de chevauchement horizontal des activités entre les deux parties à l'opération. Cependant, dans les marchés liés aux données, ce type d'opération peut se traduire par un accès différencié aux données et ainsi augmenter leur concentration si l'entrant a accès à une base de données importante (obtenue sur un autre marché par exemple)²⁸.

De même, en évaluant les risques concurrentiels inhérents à une opération de concentration, les autorités de concurrence peuvent devoir examiner de près les

²⁷ OECD, Data-Driven Innovation: Big Data for Growth and Well-Being, OECD Publishing, Paris (2015), <http://dx.doi.org/10.1787/9789264229358-en>.

²⁸ German Monopolies Commission (Monopolkommission), Special Report No. 68: Competition policy: The challenge of digital markets (2015), http://www.monopolkommission.de/images/PDF/SG/s68_fulltext_eng.pdf, §§ 109, 478; Nils-Peter Schepp and Achim Wambach, On Big Data and its Relevance for Market Power Assessment, *Journal of European Competition Law & Practice*, 2016, Vol. 7, No. 2, p. 123.

avantages pouvant être obtenus par la combinaison de différentes bases de données rendue possible par l'opération. L'association de différents ensembles de données pourrait ainsi poser des problèmes de concurrence si les concurrents de la nouvelle entité ne sont pas en mesure d'obtenir les informations que cette dernière peut obtenir en combinant ces données.

En outre, la fusion de deux entreprises qui occupent déjà des positions importantes sur des marchés amont ou aval différents peut conduire à la forclusion de ces marchés au détriment de nouveaux concurrents. Les prestataires de services en ligne consommant des volumes importants de données personnelles peuvent par exemple vouloir acquérir des producteurs d'ordinateurs, de smartphones ou de logiciels afin de s'assurer un accès durable à des quantités importantes de données par le biais des utilisateurs de ces services.

Ces considérations sont déjà prises en compte par les autorités de la concurrence et reflétées dans leur pratique décisionnelle. Dans le contexte de la fusion Facebook/WhatsApp par exemple, la Commission européenne a évalué si l'intégration entre la plateforme de réseau social Facebook et l'application de communications WhatsApp pouvait permettre à Facebook d'avoir accès à des données additionnelles auprès des utilisateurs de WhatsApp et si cela pouvait porter atteinte à la concurrence. De même, dans sa décision sur la fusion *Telefónica UK/Vodafone UK/Everything Everywhere*, la Commission a examiné « *si la JV Co évincerait les fournisseurs concurrents d'analyses de données ou de services publicitaires en associant les informations personnelles, les données de localisation, les données de réponse, les données de comportement social et les données de navigation et créerait ainsi une base de données unique qui deviendrait un intrant essentiel pour la publicité mobile ciblée qu'aucun fournisseur concurrent d'analyses de données mobiles ou de publicité aux clients ne serait en mesure de reproduire* » (§539).

Enfin, les opérations de concentration relatives aux données pourraient également générer des gains d'efficacité qui, dans la mesure où le régime de contrôle applicable aux concentrations les reconnaît, doivent être comparés aux risques qu'elles présentent pour la concurrence. Dans certains cas de concentration par exemple, les parties ont invoqué l'échelle de données à laquelle elles pourraient accéder grâce à l'opération comme relevant d'un gain d'efficacité. Dans les affaires

Microsoft/Yahoo!²⁹, United States v. Bazaarvoice et Tomtom/Tele Atlas³⁰, de tels arguments ont ainsi été formulés par les parties à l'opération en indiquant que la fusion permettrait à la nouvelle entité de produire de meilleurs produits plus rapidement grâce aux données³¹.

b) Pratiques d'éviction

Les pratiques privant certains concurrents de l'accès aux données pourraient également affaiblir la concurrence et même entraîner l'exclusion de concurrents. Différentes situations peuvent être envisagées.

aa) Refus d'accès

Le refus d'accès aux données peut être anticoncurrentiel si les données en question constituent une « facilité essentielle » à l'activité de l'entreprise qui cherche à y accéder. Toutefois, la CJUE a circonscrit l'accès obligatoire aux facilités essentielles à un nombre limité de cas, étant entendu que, même dominante, une entreprise ne peut, en principe, être obligée de favoriser l'activité de ses concurrents. Plus précisément, selon les décisions de la CJUE dans les arrêts « Bronner », « IMS Health »³² et « Microsoft »³³, une entreprise peut demander l'accès à une facilité ou à un réseau si le refus d'accès concerne un produit indispensable à l'exploitation de l'activité en question, si le refus empêche l'émergence d'un nouveau produit pour lequel il existe une demande potentielle (cette condition étant applicable lorsque l'exercice d'un droit de propriété intellectuelle est en jeu), si ce refus n'est pas justifié par des considérations objectives et s'il est susceptible d'exclure l'intégralité de la concurrence sur le marché secondaire³⁴. En outre, dans l'arrêt « Bronner », la CJUE a jugé qu'un produit ou service n'est indispensable que s'il n'existe pas de produits ou services alternatifs et si des obstacles techniques, juridiques ou économiques

²⁹ Commission européenne, "Microsoft/Yahoo!", Comp/M. 5727, en date du 18.02.2010, http://ec.europa.eu/competition/mergers/cases/decisions/M5727_20100218_20310_261202_EN.pdf, § 184.

³⁰ Commission européenne, "Tomtom/Teleatlas", Comp/M. 4854, en date du 14.05.2008, http://ec.europa.eu/competition/mergers/cases/decisions/M5727_20100218_20310_261202_EN.pdf, §§ 238-250.

³¹ Allen P. Grunes and Maurice E. Stucke, No Mistake about it: The Important Role of Antitrust in the Era of Big Data, University of Tennessee, Voir Legal Studies Research Paper No. 269 (2015), <http://ssrn.com/abstract=2600051>, p. 11.

³² ECJ, "IMS Health", C-418/01, arrêt du 29.04.2004, §§ 34-52.

³³ GC, "Microsoft", T-201/04, arrêt du 17.09.2007, §§ 320-336.

³⁴ ECJ, "IMS Health", C-418/01, arrêt du 29.04.2004, § 37.

rendent impossible ou excessivement difficile le développement de produits ou services par une entreprise, seule ou en collaboration, sur le marché en aval³⁵.

Ces critères de la CJUE ne seraient remplis que s'il est démontré que les données possédées par l'intéressé sont véritablement uniques et si son concurrent ne peut les obtenir autrement pour fournir ses services³⁶. L'amélioration de l'accès aux données peut d'ailleurs limiter les incitations que des rivaux pourraient avoir de développer leurs propres sources de données. Enfin, l'accès aux données d'une entreprise peut soulever des préoccupations en matière de protection de la vie privée si le partage forcé des données d'un utilisateur viole la législation sur la protection des données personnelles ou si les entreprises échangent des données sans avoir demandé au préalable le consentement des individus dont ils partagent les informations personnelles avec des tiers, qui n'ont pas de relations avec l'individu concerné.³⁷

bb) Accès discriminatoire aux données

Le refus d'accès aux données peut s'avérer anticoncurrentiel s'il est discriminatoire. L'affaire Cegedim en est une illustration³⁸. Cegedim, le principal fournisseur de bases de données d'informations médicales en France, a refusé de vendre sa principale base de données (appelée *OneKey*) aux clients utilisant le logiciel d'Euris, un concurrent de Cegedim sur le marché adjacent des logiciels de gestion des relations avec les clients (*customer relationship management - CRM*) dans le secteur de la santé, alors qu'elle vendait cette même base de données à d'autres clients. L'Autorité de la concurrence française a considéré que cette pratique était discriminatoire et a conclu que, étant donné que la base de données *OneKey* était la principale base de données d'informations médicales de ce type sur le marché et que Cegedim était en position dominante sur le marché des bases de données d'informations médicales, elle avait eu pour effet de limiter le développement d'Euris entre 2008 et 2012.

³⁵ ECJ, "Bronner", C-7/97, arrêt du 26.11.1998, §§ 44-45.

³⁶ Damien Geradin and Monika Kuschewsky, Competition law and personal data: preliminary thoughts on a complex issue (2013), http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2216088, p. 15.

³⁷ Darren S. Tucker and Hill B. Wellford, Big Mistakes Regarding Big Data, Antitrust Source, American Bar Association, December 2014, http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2549044, p. 11.

³⁸ Autorité française de la concurrence, Décision n° 14-D-06, en date du 08.07.2014, relative à des pratiques mises en œuvre par la société Cegedim dans le secteur des bases de données d'informations médicales. Cette décision a été confirmée en appel mais l'affaire est toujours pendante devant la Cour de Cassation.

Plus généralement, une intégration verticale peut entraîner une discrimination à l'accès à des informations stratégiques ayant pour effet des distorsions de concurrence. Ainsi, certaines places de marché agissant également comme vendeurs en ligne peuvent, par exemple, avoir accès à des informations portant sur les ventes ou sur le comportement des clients de leurs concurrents. En identifiant la gamme de produits la plus demandée, une plateforme intégrée pourrait alors être en mesure d'ajuster plus efficacement son offre et ses prix. Un effet similaire pourrait être obtenu si une telle plateforme limitait les informations que ses concurrents opérant sur la place de marché obtiennent sur les transactions dans lesquelles ils sont impliqués. Cette circulation de l'information et les limitations susceptibles de l'entraver pourraient alors conférer à l'opérateur de la plateforme intégrée un avantage concurrentiel sur ses concurrents.

cc) Contrats exclusifs

Les stratégies anticoncurrentielles basées sur les données peuvent également inclure les comportements visant à empêcher des rivaux d'accéder aux données de tierces parties au moyen d'exclusivités avec les prestataires les fournissant, ou en rendant plus difficile l'adoption par les consommateurs de leurs technologies ou l'accès à leurs plateformes³⁹. Les contrats exclusifs peuvent alors évincer des concurrents, notamment lorsque ces contrats sont conclus par des entreprises dominantes. Un réseau de contrats exclusifs peut être encore plus problématique et peut être abordé non seulement en vertu de l'Article 102 TFUE, mais également de l'Article 101 TFUE.⁴⁰ Par exemple, dans sa mise en œuvre de l'Art. 102 TFUE contre Google, la Commission européenne a examiné une série de contrats exclusifs conclus par Google dans le marché de la publicité liée aux recherches qui pourrait empêcher ses concurrents de contester sa position⁴¹.

dd) Ventes liées et utilisation croisée des ensembles de données

Les données collectées sur un marché donné peuvent également être utilisées par une entreprise pour développer ou accroître son pouvoir de marché sur un autre marché de manière anticoncurrentielle. C'est par exemple ainsi que la *UK*

³⁹ Allen P. Grunes and Maurice E. Stucke, No Mistake about it: The Important Role of Antitrust in the Era of Big Data, University of Tennessee Legal Studies Research Paper No. 269 (2015), <http://ssrn.com/abstract=2600051>, p. 3.

⁴⁰ ECJ, « Delimitis », C-234/89, arrêt du 28.02.1991 (effets cumulés d'un réseau de contrats similaires).

⁴¹ Commission européenne, « Google », affaire 38740, communiqué de presse en date du 30.11.2010, http://europa.eu/rapid/press-release_IP-10-1624_en.htm?locale=en.

Competition and Markets Authority a mentionné, dans son rapport,⁴² la possibilité de ventes liées, par lesquelles une entreprise détenant des données de valeur lie l'accès à sa base de données à l'utilisation de ses propres services d'analyse des données. De telles ventes liées peuvent alors produire des gains d'efficacité dans certains cas, mais elles peuvent également impliquer des effets restrictifs de concurrence en conférant à l'entreprise détentrice des données en question un avantage sur ses concurrents.

De manière plus générale, dans un avis de 2010, l'Autorité de la concurrence française a insisté sur le fait que l'utilisation croisée de données, c'est-à-dire l'utilisation sur un marché de données collectées sur un autre marché, peut, dans certains cas, engendrer des effets de forclusion⁴³. Notamment, les monopoles historiques ayant bénéficié d'un accès privilégié aux données dans le cadre de leur activité de service public peuvent exploiter ces données afin de proposer des offres personnalisées aux consommateurs des marchés adjacents, obtenant ainsi un avantage concurrentiel que leurs concurrents ne peuvent répliquer. Ce raisonnement a récemment conduit l'Autorité de la concurrence française à imposer des mesures conservatoire à GDF-Suez, en lui ordonnant d'accorder à ses concurrents un accès à certaines données qu'il a collectées en tant que fournisseur de gaz dans le cadre d'offres réglementées, notamment des données de consommation⁴⁴. L'objectif de cette mesure conservatoire était de permettre à tous les fournisseurs de bénéficier du même niveau d'informations pertinentes pour formuler leurs offres aux consommateurs (aucune information publique ou privée n'existant sur les ménages souscrivant des contrats de fourniture de gaz).

c) Les données comme vecteur de discrimination tarifaire

Les données sont également considérées comme pouvant faciliter la discrimination tarifaire⁴⁵. En effet, en collectant des données sur leurs clients, les entreprises peuvent obtenir des informations sur leurs habitudes d'achat et sont alors en

⁴² Competition and Markets Authority, *The Commercial Use of Consumer data* (2015), https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/435817/The_commercial_use_of_consumer_data.pdf, p.90.

⁴³ Autorité de la concurrence, Avis 10-A-13 sur l'utilisation croisée des bases de clientèle.

⁴⁴ Autorité de la concurrence, Décision 14-MC-02 du 09.09.2014. En raison des lois sur la protection des données personnelles, la transmission aux concurrents des données de GDF a été conditionnée au consentement des consommateurs. Une part importante des consommateurs a refusé que leurs données soient transmises par GDF à des opérateurs concurrents.

⁴⁵ Voir Nathan Newman, *The Costs of Lost Privacy: Consumer Harm and Rising Economic Inequality in the Age of Google*, 40 WM. MITCHELL L. REV. pp. 850 (865-873), disponible sur <http://open.wmitchell.edu/cgi/viewcontent.cgi?article=1568&context=wmlr>.

meilleure position pour évaluer leur disposition à payer pour l'achat d'un bien ou d'un service. Si cette entreprise exerce un certain pouvoir de marché⁴⁶, elle est alors en mesure d'utiliser ces informations pour fixer des prix différents selon les groupes de clients qu'elle aura identifiés grâce aux données collectées.

Différents effets de la discrimination tarifaire peuvent toutefois être envisagés. Du côté négatif, la discrimination tarifaire est souvent considérée comme une violation injuste de l'égalité entre consommateurs. De plus, si les consommateurs peuvent choisir d'acheter ailleurs pour échapper à une discrimination tarifaire qui leur serait défavorable, celle-ci peut accroître les coûts de recherche et donc réduire les possibilités de substitution. Enfin, en raison de la discrimination tarifaire, certains consommateurs finissent par payer un prix supérieur à celui en vigueur avant la mise en œuvre de la discrimination.

Pour autant, l'analyse économique montre que les effets de la discrimination tarifaire, en supposant qu'elle soit facilitée par les données, sont plus ambigus. En effet, la discrimination tarifaire basée sur les données peut certes conduire certains consommateurs à payer un bien ou un service donné à un prix supérieur, mais d'autres bénéficient en contrepartie de meilleurs prix qu'en l'absence de discrimination. Parmi ces clients, certains n'auraient d'ailleurs pas pu se permettre d'acheter le produit si un prix unique et identique avait été proposé à tous. Ces derniers sont également plus susceptibles d'être sensibles aux prix et de disposer de revenus relativement faibles⁴⁷. A cet égard, la discrimination tarifaire peut donc améliorer le bien-être social, c'est-à-dire la somme des surplus des entreprises et des consommateurs, en augmentant le nombre de transactions relativement à la situation où il n'y a pas de discrimination tarifaire. Toutefois, la maximisation du bien-être social ne se traduit pas nécessairement par celle du surplus du consommateur. En effet, la discrimination tarifaire peut aider les offreurs à fixer les prix en fonction de la disposition à payer des consommateurs de telle sorte qu'ils s'approprient la plus grande part de leur surplus.

⁴⁶ Sans aucun pouvoir de marché, un fournisseur n'est pas en mesure de fixer ses prix par référence à la propension des consommateurs à payer. En effet, un concurrent qui fixerait un prix unique par référence à ses coûts serait plus attractif que ce fournisseur vis-à-vis de la plupart des consommateurs.

⁴⁷ Executive Office of The President of The United States, Big Data And Differential Pricing (2015), https://www.whitehouse.gov/sites/default/files/docs/Big_Data_Report_Nonembargo_v2.pdf (« si les groupes défavorisés sont historiquement plus sensibles aux prix que le consommateur moyen, la tarification différentielle maximisant les profits devrait jouer en leur faveur »).

Par ailleurs, la discrimination tarifaire basée sur les données peut également avoir des effets proconcurrentiels. Ainsi, la discrimination tarifaire peut permettre à une entreprise de proposer des prix inférieurs aux clients ayant une forte préférence pour un autre produit, impliquant ainsi une plus grande concurrence tarifaire entre différents produits⁴⁸. Par conséquent, même en supposant que les données facilitent effectivement la mise en œuvre d'une discrimination tarifaire, ses effets restent toutefois difficiles à évaluer.

Enfin, on peut se demander si la discrimination tarifaire en elle-même entre dans le champ d'application du droit européen de la concurrence. Pour qu'une telle discrimination puisse être considérée comme un abus de position dominante ou une restriction verticale, ses effets sur la concurrence, et non uniquement sur le bien-être du consommateur, restent en effet à démontrer ; l'absence de justification proportionnée devrait également être établie. La situation pourrait cependant être différente sous l'empire du droit national de la concurrence. Les dispositions régissant les pratiques unilatérales peuvent en effet être plus strictes⁴⁹.

d) Données, pouvoir de marché et préoccupations en matière de vie privée

Une autre préoccupation qui a été soulevée dans la littérature explorant les questions de droit de la concurrence liées à la collecte et l'utilisation de données (personnelles) concerne les atteintes à la protection de la vie privée.

Les problématiques de confidentialité ne relèvent pas, en elles-mêmes, du champ d'intervention des autorités de concurrence. La Cour de Justice de l'Union Européenne a ainsi indiqué dans son arrêt *Asnef-Equifax* (2006) que tout aspect relatif à la sensibilité des données personnelles ne relève pas, en tant que tel, du droit de la concurrence, mais doit être traité sur la base des dispositions applicables régissant la protection des données.⁵⁰ La Commission européenne s'est inscrite dans cette approche à l'occasion de l'examen de l'opération

⁴⁸ Voir Mark Armstrong, Price discrimination, University College, London (2006), <http://else.econ.ucl.ac.uk/papers/uploaded/222.pdf>.

⁴⁹ En Allemagne, par exemple, la Cour Suprême Fédérale a indiqué que les dispositions relatives à l'abus de position dominante peuvent inclure une dimension de protection du consommateur, dans un contexte de discrimination tarifaire, Voir Cour Suprême Fédérale allemande (BGH), "Entega II", KZR 5/10, arrêt du 07.12.2010, disponible sur <http://juris.bundesgerichtshof.de/cgi-bin/rechtsprechung/list.py?Gericht=bgh&Art=en&Datum=Aktuell&Sort=12288>.

⁵⁰ ECJ, "Asnef-Equifax", C-238/05, arrêt du 23.11.2006, <http://curia.europa.eu/juris/showPdf.jsf?jsessionid=9ea7d0f130d5bea0e088b08f44b3b853d5eb7ffd88fa.e34KaxiLc3eQc40LaxqMbN4Och0Qe0?text=&docid=65421&pageIndex=0&doclang=EN&mode=lst&dir=&occ=first&part=1&cid=236269>, § 63.

« Facebook/Whatsapp », en indiquant : « *Toutes les préoccupations liées à la vie privée résultant de l'augmentation de la concentration des données relevant du contrôle de Facebook en raison de la Transaction ne tombent pas dans le champ d'application des règles de droit européen de la concurrence mais dans celui des règles européennes de protection des données* »⁵¹. Dans cette affaire, l'analyse a été centrée sur le marché des annonces publicitaires, au sein duquel les annonceurs pouvaient subir un préjudice lié à l'augmentation du pouvoir de marché de Facebook résultant de l'accroissement de ses capacités de collecte de données. L'examen n'a en revanche pas porté sur la possibilité pour Facebook de collecter davantage de données personnelles auprès des utilisateurs finaux grâce à l'opération. Ainsi qu'elle l'a expliqué, « *aux fins de la présente décision, la Commission n'a analysé la concentration potentielle des données que dans la mesure où elle est susceptible de renforcer la position de Facebook sur le marché publicitaire en ligne ou ses sous-segments* » (§164).

Pour autant, le fait que certains instruments juridiques spécifiques permettent de résoudre les risques afférents à la collecte et au traitement des données personnelles n'implique pas que l'application du droit de la concurrence soit totalement dénuée de pertinence lorsque sont en jeu des données personnelles. De manière générale, les obligations résultant d'autres corps de règles peuvent être prises en compte, au moins en tant qu'élément de contexte, dans le cadre de l'application du droit de la concurrence. Dans son arrêt *Allianz Hungária* (2013), la Cour de Justice a ainsi jugé que l'atteinte portée aux objectifs poursuivis par un autre ensemble de normes de droit national⁵² pouvait être pris en compte pour apprécier l'existence d'une restriction de concurrence (en l'espèce, caractérisée par objet)⁵³. En ce qui concerne le droit allemand de la concurrence, la Cour Fédérale de Justice a indiqué que les stipulations contractuelles qui sont incompatibles avec les lois régissant les conditions générales du commerce peuvent relever de l'abus de

⁵¹ Commission européenne, "Facebook/Whatsapp", COMP/M.7217, en date du 03.10.2014, http://ec.europa.eu/competition/mergers/cases/decisions/m7217_20141003_20310_3962132_EN.pdf, § 164.

⁵² Dans l'affaire citée, la loi hongroise exigeait que les opérateurs d'assurance agissant en tant qu'intermédiaires ou courtiers d'assurance soient indépendants des compagnies d'assurance.

⁵³ ECJ, « *Allianz Hungária* », C-32/11, jugement du 14.03.2013, <http://curia.europa.eu/juris/document/document.jsf?text=&docid=135021&pageIndex=0&doclang=EN&mode=lst&dir=&occ=first&part=1>, §§ 46, 47.

position dominante si leur mise en œuvre dépend de la dominance de l'entreprise sur le marché⁵⁴.

En effet, même si les règles en matière de protection des données et de concurrence poursuivent des objectifs différents, les politiques de confidentialité des entreprises ne peuvent, par leur seule nature, échapper à un examen au regard du droit de la concurrence. Les décisions prises par une entreprise relativement à la collecte et à l'utilisation des données personnelles peuvent en effet avoir, en parallèle des implications sur la vie privée, des implications économiques et concurrentielles. Par conséquent, les politiques de confidentialité pourraient être envisagées du point de vue du droit de la concurrence dès lors qu'elles sont susceptibles d'affecter la concurrence, notamment lorsqu'elles sont mises en œuvre par une entreprise dominante pour laquelle les données servent d'intrant principal pour ses produits ou ses services. Un lien étroit entre la dominance de l'entreprise, ses processus de collecte de données et la concurrence qui prévaut sur les marchés concernés peut alors être établi, justifiant ainsi la prise en considération des politiques et des règles de confidentialité dans des procédures de concurrence.

Il peut également être intéressant de considérer les aspects relatifs à la diversité culturelle. La diversité culturelle a en effet été explicitement prise en compte par la Commission européenne, notamment dans le domaine du contrôle des fusions (voir la décision Universal/EMI). L'Article 167(4) du Traité sur le fonctionnement de l'Union européenne dispose que « *L'Union tient compte des aspects culturels dans son action au titre d'autres dispositions des traités, afin notamment de respecter et de promouvoir la diversité de ses cultures* ». L'Article 16 TFUE, bien qu'il n'impose pas explicitement que les aspects de protection des données soient considérés dans toutes les actions menées par l'Union en vertu des Traités, affirme cependant que « *Toute personne a droit à la protection des données à caractère personnel la concernant* ». Même si la question de savoir si ces dispositions emportent des obligations spécifiques pour les autorités de concurrence reste ouverte, le *Contrôleur européen de la protection des données* a recommandé en 2014 une évolution de la politique de concurrence et une « *approche plus holistique de sa mise en oeuvre* » dans laquelle un dialogue plus systématique serait maintenu entre les autorités

⁵⁴ Cour Fédérale de Justice allemande (Bundesgerichtshof), "VBL-Gegenwert", KZR 61/11, jugement du 16.11.2013, disponible sur <http://juris.bundesgerichtshof.de/cgi-bin/rechtsprechung/list.py?Gericht=bgh&Art=en&Datum=Aktuell&Sort=12288>, § 68.

chargées de la concurrence, de la protection des données et des consommateurs « dès que se présente un cas spécifique dans lequel des préoccupations concernant le bien-être du consommateur et la protection des données semblent être en jeu »⁵⁵.

Dans les cas de contrôle des concentrations, la question de la protection des données pourrait devenir particulièrement pertinente du point de vue de la concurrence si une entreprise donnée devait bénéficier d'un fort pouvoir de marché vis-à-vis de ses utilisateurs finaux. En effet, les entreprises qui obtiennent une position de marché avantageuse en raison d'une fusion peuvent être en mesure d'en renforcer les effets via une collecte de données plus importantes concernant les consommateurs ainsi qu'une dégradation de la politique de confidentialité. Si deux entreprises sont en concurrence sur le terrain de la protection de la vie privée, en tant que facteur de différenciation de qualité, leur fusion peut alors réduire la qualité⁵⁶. Pour certains analystes⁵⁷, la réduction de la confidentialité est en effet synonyme de réduction de qualité. A ce jour, un tel scénario n'a pas connu d'application puisque la protection de la vie privée n'est pas encore apparue comme un paramètre concurrentiel significatif dans la pratique décisionnelle des autorités de concurrence.

La dégradation de la politique de confidentialité pourrait également être appréhendée sous l'angle de l'abus de position dominante si un opérateur collecte des données en violant de manière manifeste la législation relative à la protection des données personnelles, et s'il existe une forte interaction entre la collecte des données et la position sur le marché de l'entreprise. Jusqu'à présent, les autorités de concurrence ont assimilé en grande partie les abus d'exploitation aux pratiques de prix excessifs⁵⁸. Ces interventions sont l'objet de nombreuses difficultés pratiques car elles nécessitent de trouver un marché comparable, ou de pouvoir opérer des comparaisons complexes entre prix et coûts ainsi que de définir des prix « de référence ». Certains avancent que ces difficultés pratiques, et le risque qu'elles

⁵⁵ European Data Protection Supervisor, Privacy and competitiveness in the age of big data, preliminary opinion (2014), https://secure.edps.europa.eu/EDPSWEB/webdav/site/mySite/shared/Documents/Consultation/Opinions/2014/14-03-26_competition_law_big_data_EN.pdf.

⁵⁶ CMA, The commercial use of consumer data (2015), https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/435817/The_commercial_use_of_consumer_data.pdf, p. 95.

⁵⁷ Allen P. Grunes and Maurice E. Stucke, No Mistake about it: The Important Role of Antitrust in the Era of Big Data, University of Tennessee. Voir Legal Studies Research Paper No. 269 (2015), <http://ssrn.com/abstract=2600051>, p. 4.

⁵⁸ Emil Paulis, "Art. 82 EC and Exploitative Conduct", in Ehlermann/Marquis, European Competition Law Annual 2007: A Reformed Approach to Article 82 EC, p. 2.

conduisent les autorités de concurrence à adopter des solutions sous-optimales, sont si importants que les actions à l'encontre des abus d'exploitation ne devraient être entreprises qu'en dernier recours⁵⁹. Cependant, à l'occasion de l'examen de certaines conditions commerciales « excessives », notamment les conditions générales qui sont imposées aux consommateurs en vue d'utiliser un service ou un produit, les règles en matière de protection des données personnelles pourraient constituer une référence utile pour évaluer une pratique abusive, notamment dans un contexte où la plupart des consommateurs ne lisent pas les conditions générales de vente ni les politiques de confidentialité des différents prestataires de services qu'ils utilisent⁶⁰.

IV. Big data et pouvoir de marché

Parmi les effets que la collecte et l'utilisation des données peuvent avoir, et qui ont été précédemment mentionnés, la question de la possible contribution des données au pouvoir de marché est la principale et elle est, à de nombreux égards, la plus intéressante du point de vue de l'analyse concurrentielle. Plusieurs des scénarios anticoncurrentiels présentés dans la partie III dépendent de la mesure dans laquelle les données peuvent conférer un avantage aux entreprises, en particulier à celles qui sont mieux positionnées pour procéder à leur collecte massive. Cette question est complexe. S'agissant tout particulièrement des activités économiques en ligne, la question des données n'épuise pas les aspects quelque peu « non traditionnels » à considérer dans l'évaluation du pouvoir de marché des entreprises. Cela doit être gardé à l'esprit (1). Ensuite, afin d'évaluer si les données contribuent effectivement à la création ou à la préservation d'un pouvoir de marché, les autorités de concurrence devront évaluer l'ampleur de l'avantage économique que les données confèrent. C'est un exercice qui a déjà été mis en œuvre dans différents cas, même s'il n'a trouvé à s'appliquer, s'agissant des données personnelles, que dans le cadre du contrôle des concentrations sur des marchés numériques (2). Ces décisions, ainsi que plusieurs contributions académiques ou de consultants, soulignent l'attention qui

⁵⁹ Emil Paulis, "Art. 82 EC and Exploitative Conduct", in Ehlermann/Marquis, *European Competition Law Annual 2007: A Reformed Approach to Article 82 EC*, p. 3.

⁶⁰ L'eurobaromètre de la protection des données (juin 2015) montre qu'un cinquième seulement des personnes interrogées ont intégralement lu les avertissements sur la confidentialité (18%) http://ec.europa.eu/justice/data-protection/files/factsheets/factsheet_data_protection_eurobarometer_240615_en.pdf; Le Bundeskartellamt a récemment initié une procédure à l'encontre de Facebook pour enquêter sur des allégations que Facebook a abusé d'une position possiblement dominante sur le marché des réseaux sociaux par ses conditions spécifiques d'utilisation des données de l'utilisateur.

devrait être portée à la disponibilité des données sur les marchés numériques (3) et aux avantages spécifiques qui découlent d'un accès plus massif aux données (4).

1. Les sources de pouvoir de marché dans les activités en ligne consommatrices de données

De nombreux secteurs, y compris ceux reposant pour l'essentiel sur la vente de biens matériels par le biais de « circuits en dur », pourraient être concernés par les possibilités actuelles et futures de collecte et d'utilisation des données. Toutefois, l'essentiel du débat sur la manière dont ces possibilités pourraient affecter la concurrence s'est, jusqu'à présent, centré sur les services en ligne et plus spécifiquement sur les moteurs de recherche, les réseaux sociaux et la distribution de détail. En effet, grâce à leur taille, aux circuits numériques qu'elles empruntent ainsi qu'aux appareils qui sous-tendent leurs interactions avec le marché, les entreprises en ligne collectent souvent plus de données que leurs homologues en dur. Les entreprises en ligne peuvent également être mieux à même d'utiliser ces données dans des délais rapides, grâce à des processus automatisés, par exemple à des fins publicitaires ou de recommandation. En outre, ces données peuvent fonder de nombreux modèles économiques en ligne, notamment ceux qui sont financés par la publicité. Enfin, certains marchés en ligne sont perçus par différents observateurs comme excessivement concentrés, engendrant la crainte que le développement de la collecte et de l'exploitation des données n'aggrave cette concentration.

Cela fait plus d'une décennie que les entreprises en ligne et leurs caractéristiques économiques font l'objet d'âpres discussions entre universitaires⁶¹. La littérature économique identifie ainsi certains aspects qui caractérisent plusieurs activités en ligne et qui doivent être pris en considération dans l'évaluation du pouvoir de marché des entreprises. Ces aspects incluent notamment les effets de réseau, le multi-hébergement (multi-homing) et une dynamique de marché spécifique.

Le débat académique sur ces facteurs est toujours en cours. La pratique décisionnelle a cependant montré que de tels phénomènes pouvaient avoir des

⁶¹ David S. Evans and Richard SchmalenVoir, *The industrial Organization of Markets with Two-sided Platforms*, Competition Policy International, 2007, Vol. 3, p. 164; Bernard Caillaud and Bruno Jullien, *Chicken & egg: competition among intermediation service providers*, RAND Journal of Economics 2003, 34(2), pp. 309-328; Mark Armstrong, *Competition in two-sided markets*, RAND Journal of Economics 2006, 37(3), pp. 668-691; Lapo Filistrucchi/Damien Geradin/Eric van Damme, [Identifying two-sided markets](#), World Competition: Law and Economics Review 2013, 36 (1), pp. 33-60.

effets ambigus sur la concurrence et qu'aucune conclusion définitive et de portée générale ne pouvait être tirée, ce qui appelle une analyse au cas par cas. Par conséquent, l'évaluation concurrentielle doit être extrêmement fine et se rattacher au cas d'espèce.

a) Marchés multi-faces et effets de réseau

De nombreux marchés en ligne sont dits « multi-faces » lorsque les entreprises y opérant s'adressent à plusieurs groupes d'utilisateurs/de consommateurs. Cette particularité peut être à considérer pour la définition des marchés pertinents. En outre, certains services sont fournis sans contrepartie financière sur une face du marché, ce qui pose la question de savoir si cela constitue un obstacle à la définition d'un marché pertinent (ce qui a été par exemple le cas en Allemagne) en dépit du fait que les entreprises présentes du « côté gratuit » du marché sont en concurrence sur d'autres aspects que le prix, comme la qualité des services ou - vraisemblablement dans une moindre mesure jusqu'à présent - la quantité de données collectées sur chaque utilisateur⁶².

En outre, des « effets de réseau » sont souvent à l'œuvre dans ce type de cas. Le terme « effets de réseau » se réfère à la manière dont l'utilisation d'un bien ou d'un service par un utilisateur donné influence la valeur que ce produit a pour d'autres utilisateurs. Ces effets peuvent être « directs » lorsque l'avantage obtenu par les utilisateurs d'un groupe dépend du nombre d'autres utilisateurs de ce même groupe qui utilisent ce service. Les réseaux de télécommunication en sont l'exemple classique : plus les gens les utilisent et peuvent être joints par leur biais, plus ils leur apportent de l'utilité. Les effets de réseau peuvent également être « indirects », lorsque le bénéfice que les utilisateurs d'un groupe tirent du service dépend du nombre des utilisateurs de ce service qui appartiennent à un groupe différent du leur. Une plateforme de rencontre permettant à des hommes et des femmes de se rencontrer en est une illustration. Les effets de réseaux directs et indirects peuvent également coexister dans certains cas. Par exemple, la valeur d'un réseau social pour un utilisateur donné est susceptible de croître avec le nombre total des utilisateurs de ce réseau (effets de réseau directs) tandis que, dans le même temps, un nombre supérieur d'utilisateurs d'un réseau social augmente également sa valeur pour les annonceurs (effets de réseau indirects). Il doit toutefois être souligné que les

⁶² Voir GC, « Microsoft », T-201/04, jugement du 17.09.2007, §§ 966-970; GC, « Cisco Systems », T-79/12, jugement du 11.12.2013, §§ 65-74.

effets de réseau indirects ne sont pas nécessairement symétriques. Par exemple, la valeur d'un réseau social pour les annonceurs augmente avec le nombre d'utilisateurs mais il est incertain, voire douteux, que les utilisateurs attachent de la valeur à un nombre supérieur d'annonceurs ou de publicités, bien qu'ils puissent en accorder aux différents investissements effectués par le réseau social pour améliorer ses services grâce aux revenus publicitaires. Ainsi, la portée des effets de réseau doit être évaluée au cas par cas car leur importance absolue, comme la manière dont ils évoluent avec la conquête de nouveaux utilisateurs, peuvent varier en fonction du service considéré.

Les effets de réseau peuvent avoir des conséquences contrastées sur la concurrence. Souvent, ils sont cités avec à l'esprit un effet « boule de neige » qui peut favoriser la concentration du marché. Ils sont également considérés comme un obstacle potentiel à l'entrée et, partant, comme un facteur de limitation de la concurrence. Dans ce contexte, la collecte et l'exploitation de données peut renforcer les effets de réseau lorsque l'accroissement du nombre d'utilisateurs d'une entreprise lui permet de collecter davantage de données que ses concurrents, ce qui lui permet ensuite d'accroître la qualité de ses produits ou services et au final d'accroître ses parts de marché (voir partie III ci-dessus).

A contrario, les effets de réseau peuvent également être bénéfiques à de nouveaux entrants sur le marché s'ils sont en mesure d'attirer un nombre élevé d'utilisateurs pour d'autres raisons que leur taille (grâce par exemple à une fonctionnalité innovante), augmentant ainsi leur attractivité auprès d'autres utilisateurs grâce aux effets de réseau. Les effets de réseau peuvent donc également stimuler la concurrence en donnant à un nouvel entrant la capacité d'augmenter rapidement sa base d'utilisateurs. En fonction de différents paramètres, tels que le niveau des coûts fixes ou les différences de parts de marché des entreprises, les effets de réseau peuvent ainsi renforcer ou, au contraire, atténuer, la concurrence.

b) Multi-hébergement (« multi-homing »)

Des consommateurs sont dits multi-hébergés lorsqu'ils utilisent plusieurs prestataires pour obtenir un même type de service.

De nombreux économistes considèrent que le multi-hébergement est un facteur susceptible de réduire le pouvoir de marché des entreprises⁶³. Mais une situation de multi-hébergement parfait, dans laquelle chaque utilisateur d'un service fait appel à plusieurs offreurs, est plutôt rare, en raison de coûts de transfert de nature variée (effets de réseau, coûts d'apprentissage, etc.). En outre, la collecte des données peut également augmenter ces coûts de transfert car le prestataire le plus utilisé par un consommateur donné détiendra plus d'informations sur lui et, par conséquent, sera en situation de mieux le satisfaire en personnalisant son offre de services. En pratique, une situation intermédiaire entre l'hébergement unique et le multi-hébergement est la plus courante. En outre, les interactions entre le multi-hébergement (dans une certaine mesure) sur l'une des faces d'un marché biface et l'hébergement unique (dans une certaine mesure) sur l'autre face de ce même marché doivent être prises en compte pour analyser les effets dans un cas donné.

Il est également souvent avancé que les marchés où les services sont gratuits pour les utilisateurs finaux, sont propices au multi-hébergement. Cette affirmation doit cependant être nuancée. Ainsi, des coûts de transfert peuvent empêcher les consommateurs d'utiliser différents prestataires avec la même fréquence. Les consommateurs peuvent en outre accorder une attention plus importante à la qualité lorsque les services sont gratuits, car dans ce contexte, il s'agit du seul facteur de concurrence entre les plateformes. Or, dans un marché caractérisé par des effets de réseau et d'expérience (pouvant être liés aux volumes de données collectées)⁶⁴, les nouveaux entrants peuvent ne pas être en mesure de proposer une qualité de service comparable à celle d'entreprises déjà établies et ne pourraient compenser cette moindre qualité par des prix inférieurs.

Enfin, alors que la possibilité de multi-hébergement - et, plus généralement, les faibles coûts de transfert - peut limiter le pouvoir de marché des entreprises en place (en rendant plus aisée la substitution si la qualité des services fournis par ces entreprises diminue), elle n'emporte pas nécessairement de conséquence lorsqu'on considère le pouvoir de marché des entreprises basé sur les données. En effet, la valeur ajoutée des données d'utilisateurs ne peut faire une différence pour une

⁶³ David S. Evans and Richard Schmalen, Voir *The Industrial Organisation of Markets with Two-Sided Platforms*, Competition Policy International, 2007, Vol. 3, Nr. 1, pp. 151-179; Mark Armstrong, *Competition in two-sided markets*, RAND Journal of Economics, Vol. 37, Nr. 3, pp. 668-691.

⁶⁴ Les effets d'expérience désignent la réduction des coûts (ou, dans certains cas, les améliorations de la qualité) des produits à mesure que leurs volumes de production/consommation augmentent.

entreprise qu'à la condition que ces utilisateurs ne recourent au multi-hébergement de leurs données, et n'utilisent donc des prestataires rivaux, que de manière peu fréquente ou très ponctuelle. Or un multi-hébergement effectif peut s'avérer illusoire lorsque des effets du réseau et d'expérience entrent en jeu.

c) Dynamique de marché

Un autre argument souvent avancé consiste à relativiser le pouvoir de marché sur les marchés numériques en raison du déploiement rapide de produits innovants. Un « *nombre considérable d'entrées et de gains rapides souvent obtenus par de nouveaux concurrents innovants* »⁶⁵ sur ces marchés pourrait alors témoigner du fait que les obstacles à l'entrée et au développement, qu'ils soient liés aux besoins de données ou à d'autres types d'actifs, sont finalement peu importants. Les partisans de cet argument de la « concurrence dynamique » se réfèrent souvent au remplacement de Yahoo!, Lycos ou AltaVista par Google et de MySpace par Facebook en quelques années seulement. Or, lorsqu'ils sont entrés sur le marché, Google et Facebook ne bénéficiaient pas des mêmes effets de réseau et d'expérience et des mêmes économies d'échelle que leurs rivaux. De même, ils n'avaient pas accès aux mêmes volumes de données que les leaders du marché mais ils ont pourtant surmonté ce handicap. Dans la décision Facebook/Whatsapp, au stade de l'analyse du marché des applications de communication grand public, la Commission a estimé que « *les applications de communications grand public sont un secteur en évolution rapide, dans lequel les coûts de transfert et les obstacles à l'entrée ou à l'expansion sont faibles. Dans ce marché, il est peu probable que l'entreprise leader, même si elle est soutenue par des effets de réseau, ne puisse être contestée. Le marché des applications de communications grand public a une longue histoire d'entrées par de nouveaux acteurs. Les applications de communications en concurrence sont également en mesure de croître en dépit des effets de réseau, tant sur la durée qu'à la suite de perturbations du marché. La menace présentée par les nouveaux acteurs constitue, et continuera probablement à constituer un facteur significatif de discipline pour l'entité issue de la concentration, indépendamment de la taille de son réseau* » (§ 132).

⁶⁵ Darren S. Tucker and Hill B. Wellford, Big Mistakes Regarding Big Data, Antitrust Source, American Bar Association, December 2014, <http://ssrn.com/abstract=2549044>.

Bien qu'une telle dynamique concurrentielle puisse contrebalancer des préoccupations liées à un pouvoir de marché statique, au moins sur certains marchés, cet effet doit être évalué au cas par cas, et les difficultés d'entrée et de développement des nouvelles entreprises sur les marchés en ligne ou numériques ne doivent pas être sous-estimées et nécessitent un examen attentif, notamment au regard du rôle que la collecte de données peut avoir dans cette dynamique. Selon le cas considéré, les coûts d'entrée sur le marché peuvent inclure les dépenses de recherche et développement, les actifs matériels nécessaires à l'exploitation, les dépenses de marketing pour faire connaître le service à ses utilisateurs potentiels, etc. Les entrées sur les marchés en ligne peuvent également requérir un examen plus attentif. Nombreux sont les nouveaux entrants sur ces marchés à se concentrer sur des niches, qu'il s'agisse de fonctionnalités particulières, de segments de clientèle, ou de domaines d'intérêt des utilisateurs. Certains de ces nouveaux entrants prometteurs, notamment des noms bien connus d'Internet, n'ont toutefois pas tenu leurs promesses, l'exemple le plus connu étant peut-être celui du réseau social « Google+ ». Enfin, des exemples passés d'entrants ayant été couronnés de succès, comme les cas de Google ou Facebook sur leurs marchés respectifs, ne reflètent pas nécessairement les conditions d'entrée prévalant actuellement, celles-ci ayant changé depuis leur entrée. Ainsi, l'accroissement du rôle joué actuellement par les données dans le développement de nouveaux services reste à évaluer. Enfin, la dynamique concurrentielle peut être ralentie si les entreprises en place sont en mesure d'acquiescer de nouveaux entrants susceptibles d'animer la concurrence sans, pour autant, eu égard à leur chiffre d'affaires, leur base d'actifs ou leur part de marché, déclencher d'obligation de notification de l'opération au titre du contrôle des concentrations.

2. Évaluation de l'avantage dérivé des données dans des affaires passées

a) Sur les marchés « non numériques »

L'évaluation des avantages conférés par certaines bases de données à une entreprise est un exercice que les autorités de concurrence ont déjà pratiqué dans plusieurs affaires.

Dans son avis du 14 juin 2010 sur l'utilisation croisée des bases de clientèle⁶⁶, c'est-à-dire l'utilisation par une entreprise de bases de données sur les clients collectées sur un marché afin de développer son activité sur un autre marché, l'Autorité de la concurrence française a expliqué que les critères utilisés pour déterminer si l'utilisation de telles bases de données était susceptible de restreindre la concurrence incluent les conditions dans lesquelles elles ont été constituées, si elles peuvent être reproduites par les concurrents dans des conditions raisonnables et, enfin, si leur utilisation est susceptible de se traduire par un avantage concurrentiel significatif.

L'Autorité de la concurrence française a appliqué ce raisonnement dans différentes affaires, impliquant notamment d'anciens monopoles publics. Dans la décision précédemment mentionnée de 2014 relative à GDF, l'Autorité de la concurrence a ainsi observé que les informations contenues dans la base de données clientèle collectées par GDF incluaient des coordonnées, des informations techniques et des données de consommation. Or, cette base de données était utilisée par l'entreprise pour proposer des offres personnalisées à tous ses clients, passés et actuels, ce qui l'a aidé à vendre ses services sur le marché du gaz. L'Autorité de la concurrence a estimé que la base en question ne résultait pas d'une innovation particulière de GDF, puisqu'elle était héritée de son monopole historique, et qu'elle ne pouvait pas être reproduite dans des conditions financières raisonnables et dans un délai suffisamment court par ses concurrents. Dans ce contexte, en dépit du fait que ces données n'étaient pas considérées comme un actif essentiel, l'Autorité de la concurrence a conclu qu'elles donnaient à GDF un avantage concurrentiel significatif injustifié et que GDF était ainsi susceptible d'avoir abusé de sa position dominante sur le marché du gaz en les utilisant⁶⁷.

Ce type de raisonnement et d'évaluation de l'avantage ainsi dérivé de la détention de données se retrouve dans d'autres affaires. L'Autorité de la concurrence⁶⁸ a ainsi sanctionné EDF pour avoir abusé de sa position dominante sur le marché de la fourniture d'électricité en utilisant des données collectées auprès de ses abonnés sur

⁶⁶ Autorité de la concurrence, Avis n°10-A-13 du 14.06.2010, <http://www.autoritedelaconcurrence.fr/pdf/avis/10a13.pdf>. Cet avis de l'Autorité visait à donner des directives générales sur le sujet. Il ne concernait pas un marché ou un secteur donné mais décrivait une application possible de l'analyse à l'industrie des télécommunications.

⁶⁷ Cette décision de mesures conservatoires a été confirmée en appel.

⁶⁸ Autorité de la concurrence, décision n°13-D-20 of 17.12.2013, confirmée sur ces points par la Cour d'appel le 21.05.2015.

le marché de l'électricité (notamment des coordonnées) afin de promouvoir efficacement et à faible coût les offres de l'une de ses filiales sur un autre marché, celui de la production d'électricité photovoltaïque. Cette base de données, en partie héritée de son monopole légal sur la fourniture d'électricité à des prix réglementés, n'était pas accessible aux concurrents et a été considérée comme non reproductible dans des conditions financières et dans un délai raisonnable, notamment en raison de sa taille (20 millions de clients).

Un raisonnement similaire a également été tenu dans des cas de concentration. Ainsi, dans sa décision sur la fusion EDF-Dalkia⁶⁹, la Commission européenne a examiné la manière dont les informations sur la consommation d'électricité pouvaient affecter la concurrence sur le marché de la fourniture de services permettant l'optimisation de la consommation d'électricité. Elle a observé que l'accès à ces informations n'était pas susceptible de donner à la nouvelle entité un avantage concurrentiel significatif étant donné que les clients (des entreprises et des coopératives qui se procurent ces services par le biais d'un processus d'appel d'offres) pouvaient recevoir ces informations du gestionnaire du réseau (ERDF) puis les fournir à tout fournisseur concurrent d'EDF-Dalkia. Dans la décision portant sur la fusion Enerest/Electricité de Strasbourg⁷⁰, l'Autorité de la concurrence était préoccupée par le risque que la nouvelle entité, qui résulterait du rapprochement des deux fournisseurs historiques de gaz et d'électricité dans la région de Strasbourg, soit le seul acteur du marché ayant un accès exclusif à des données complètes de consommation d'électricité et de gaz obtenues au travers des monopoles légaux respectifs de fourniture du gaz et de l'électricité au tarif réglementé dont jouissaient les deux entités. Ces données lui permettraient en effet de proposer des offres combinées de gaz et d'électricité spécifiquement adaptées à chaque client potentiel, donnant ainsi à la nouvelle entité un avantage concurrentiel. Afin de disposer du même niveau d'informations, les concurrents devraient alors adresser des demandes à leurs prospects, impliquant un coût que l'entité issue de la concentration n'aurait pas à engager⁷¹. L'Autorité de la concurrence a autorisé l'opération parce que Enerest et Electricité de Strasbourg se sont finalement engagées à adresser à

⁶⁹ Commission européenne, « EDF/Dalkia en France », COMP/M.7137, en date du 25.06.2014, http://ec.europa.eu/competition/mergers/cases/decisions/m7137_20140625_20310_3804903_FR.pdf

⁷⁰ Autorité française de la concurrence, décision n°12-DCC-20 du 07.02.2012.

⁷¹ Autorité française de la concurrence, décision n°12-DCC-20 du 07.02.2012, § 71.

chaque concurrent qui le demanderait les informations nécessaires à la conception d'offres commerciales combinées et personnalisées (§§ 79-80, §§ 87-90).

Le contexte de la décision Cegedim mentionnée *supra* est différent dans la mesure où Cegedim, bien qu'en position dominante sur le marché de la fourniture de bases de données médicales à des sociétés pharmaceutiques et à des visiteurs médicaux, n'était pas un ancien monopole public. En outre, le scénario anticoncurrentiel ne reposait pas sur l'utilisation exclusive des données par Cegedim. L'Autorité de la concurrence a cependant mis l'accent, comme dans les autres décisions précédemment mentionnées, sur la haute qualité de la base d'informations médicales OneKey conçue et vendue par Cegedim, en termes de couverture et de mises à jour, ainsi que sur les coûts élevés de reproduction d'une telle base de données. Bien que la base OneKey n'était pas considérée comme une facilité essentielle pour les concurrents, comme Euris, sur le marché aval des logiciels de gestion de la relation clients (qui employait des bases de données du même type que OneKey), l'Autorité de la concurrence a conclu que les conditions d'accès à cette base, qui étaient discriminatoires pour les clients d'Euris, entravaient de manière significative la concurrence en détériorant la qualité de ses services et son image commerciale relativement à ses concurrents.

b) Sur les marchés « numériques »

Bien que le rôle des données sur les marchés non numériques puisse être très important dans certains cas, et qu'il soit susceptible d'augmenter à l'avenir à mesure que les objets connectés deviennent omniprésents, c'est véritablement s'agissant des marchés numériques que la collecte et l'utilisation de vastes volumes de données ont attiré le plus d'attention. Les parties suivantes décrivent certains des critères utilisés pour évaluer la manière dont les données peuvent avoir de l'influence sur la concurrence et, partant, sur l'analyse concurrentielle de ces marchés. Il existe ainsi plusieurs exemples, exclusivement dans le contexte du contrôle des concentrations, où les autorités de concurrence ont examiné les avantages que la collecte de données sur les marchés numériques peut conférer aux entreprises. Dans l'ensemble de ces cas, les autorités de concurrence impliquées, c'est à dire les autorités américaines et la Commission européenne, ont considéré que l'avantage que la nouvelle entité pourrait tirer des données n'entraînait pas de risque pour la concurrence sur les marchés considérés.

Plus précisément, comme indiqué dans la partie II, les données sont souvent utilisées dans les marchés numériques pour améliorer la qualité du service offert aux clients. Les moteurs de recherche les plus connus utilisent ainsi les données collectées à l'occasion des recherches effectuées par leurs utilisateurs pour améliorer la qualité des résultats proposés lors de recherches futures. Les boutiques en ligne utilisent également les données concernant les achats effectués ou les transactions non abouties pour recommander des produits à leurs clients. Les réseaux sociaux quant à eux sélectionnent les informations les plus pertinentes à proposer à un utilisateur donné en se basant sur l'analyse de l'activité de cet utilisateur sur le réseau social. Enfin, l'accès aux données d'un utilisateur permet à une plateforme de cibler les annonces publicitaires qu'elle émet en fonction des caractéristiques individuelles de ses utilisateurs ou des caractéristiques d'un groupe d'utilisateurs.

La manière dont ces processus de collecte de données permettent à leurs propriétaires de fournir de meilleurs services à leurs clients est bien identifiée par ces décisions de contrôle des concentrations. Dans sa décision Google/DoubleClick par exemple, la Commission a estimé que *« contrairement à la publicité hors ligne, la publicité en ligne est considérée comme capable d'atteindre un public plus ciblé de manière plus efficace. Les annonceurs peuvent cibler précisément leur public en combinant les informations relatives à leur situation géographique, à l'heure de la journée, aux domaines d'intérêt, aux précédents achats de l'utilisateur et aux préférences de recherche »* (§ 45, soulignement ajouté). Ainsi, dès 2008, la Commission reconnaissait déjà le rôle que les données peuvent jouer en aidant les annonceurs, ainsi que leurs intermédiaires, à personnaliser leurs annonces en fonction des caractéristiques de l'utilisateur auxquelles elles s'adressent. La Commission a analysé la possibilité que la nouvelle entité cherche à combiner les données collectées par DoubleClick puis par Google, ce qui pourrait lui permettre d'accroître la qualité de ses services : *« Une telle combinaison utilisant les informations sur les adresses IP des utilisateurs, les identifiants de cookies et les horaires de connexion pour rapprocher correctement les historiques des deux bases de données, pourrait impliquer que les historiques de recherches des différents utilisateurs soient liés au comportement de navigation de ces mêmes utilisateurs sur internet. Par exemple, à la suite d'un tel rapprochement, l'entité issue de la concentration pourrait savoir que le même utilisateur a recherché les termes A, B et*

C et visité les pages web X, Y et Z au cours de la semaine précédente. Ces informations pourraient potentiellement être utilisées pour mieux cibler les annonces aux utilisateurs. »⁷². La Commission a cependant estimé qu'une telle combinaison pourrait être difficile à mettre en œuvre en pratique, notamment en raison de la potentielle réticence des annonceurs (§§ 361-363), et que, même si elle pouvait être mise en place, elle ne conférerait probablement pas à la nouvelle entité un avantage concurrentiel irrattrapable par ses concurrents. En effet, plusieurs d'entre eux exploitent un moteur de recherche tout en proposant également la publication d'annonces, comme le ferait l'entité issue de l'opération. En outre, la Commission a jugé que les concurrents pouvaient acquérir des données ou des services de ciblage auprès de tiers, ce qui pouvait compenser un moindre accès aux données par rapport à l'entité issue de l'opération (§§ 364-366). La possibilité que les concurrents accèdent à des données similaires a par conséquent été l'un des principaux critères pris en compte par la Commission (voir ci-après).

De même, dans sa décision Facebook/Whatsapp, la Commission a évoqué la possibilité que « *l'entité issue de la concentration puisse commencer à collecter les données des utilisateurs Whatsapp dans le but d'améliorer le ciblage des annonces diffusées sur la plateforme de réseau social de Facebook* »⁷³. Elle a observé cependant que les incitations de l'entité issue de l'opération étaient incertaines puisqu'une telle collecte de données pourrait pousser certains utilisateurs à se tourner vers d'autres applications de communication jugées moins intrusives (§ 186), et que, en tout état de cause, la quantité de données individuelles pouvant être utilisées à des fins publicitaires et qui ne relèvent pas du contrôle exclusif de Facebook restait très vaste (§ 189). Dans ce contexte, la Commission a estimé que la fusion ne donnait pas lieu à des préoccupations de concurrence sur le marché des services publicitaires en ligne.

Dans sa décision sur la fusion Microsoft/Yahoo! Search Business, la Commission est allée un peu plus loin en considérant que l'accès aux données (comme par exemple, les données portant sur les recherches effectuées par les utilisateurs ou les liens sur lesquels ils cliquaient) était déterminante pour la performance d'un moteur de

⁷² Commission européenne, "Google/DoubleClick", COMP/M. 4731, en date du 11.03.2008, http://ec.europa.eu/competition/mergers/cases/decisions/m4731_20080311_20682_en.pdf, § 360.

⁷³ Commission européenne, "Facebook/WhatsApp", COMP/M.7217, en date du 03.10.2014, http://ec.europa.eu/competition/mergers/cases/decisions/m7217_20141003_20310_3962132_EN.pdf, § 180.

recherche mais, également, que l'accroissement des possibilités de collecte de données permettrait à la nouvelle entité d'améliorer sa capacité à fournir aux utilisateurs de meilleurs services. Ainsi, la Commission a indiqué qu' « *il est plausible que l'entité issue de la concentration soit, par l'innovation et son accès à un répertoire plus large, en mesure de fournir des résultats de recherche personnalisés mieux alignés sur les préférences des utilisateurs* » (§§ 225-226)⁷⁴. La Commission a ensuite considéré que l'opération ne présentait pas de risque pour la concurrence et qu'elle était même susceptible d'avoir un impact positif sur la concurrence, étant donné – en résumé – que les moteurs de recherche de Yahoo ! et de Microsoft étaient tous deux à la traîne par rapport à celui de Google, de telle sorte qu'un accès plus large aux données contribuerait à renforcer la pression concurrentielle que l'entité issue de l'opération pourrait exercer sur Google.

Ces décisions de concentration relatives aux marchés numériques permettent d'identifier deux facteurs susceptibles d'être particulièrement pertinents pour déterminer si les données peuvent ou non conférer du pouvoir de marché, à savoir la rareté des données (ou la facilité avec laquelle elles peuvent être reproduites) et l'influence éventuelle du volume ou de la variété des données collectées sur la compétitivité des entreprises.

3. Aspects liés à la rareté des données

Les données peuvent influencer sur la concurrence si elles fournissent à leurs propriétaires un avantage concurrentiel. Si les données sont largement disponibles, aucune entreprise, même dominante, ne jouira d'avantages comparatifs spécifiquement liés à la détention de ces données, ses concurrents étant tout autant en mesure d'obtenir les informations et les gains d'efficacité dérivés de ces données.

Il existe en particulier trois facteurs susceptibles de faire que les données soient largement disponibles et de réduire ainsi le risque qu'un accès différencié à ces données puisse affecter la concurrence. Tout d'abord, les données sont, par définition, « non rivales » de sorte que l'accès aux données par un opérateur ne prive pas nécessairement les autres de la possibilité d'accéder à ces mêmes données. Deuxièmement, les courtiers en données se sont développés, en particulier aux Etats-Unis, et pourraient contribuer à une plus grande disponibilité des données.

⁷⁴ La Commission a également indiqué que les retours sur une base plus large décroitraient probablement. Voir ci-dessous.

Enfin, les marchés numériques sont particulièrement propices à la collecte de données, ce qui peut encore augmenter la disponibilité des données. Ces trois aspects sont envisagés tour à tour.

a) Sur le caractère non rival des données

aa) Présentation

Les données sont non rivales, ce qui implique que la détention et l'utilisation par une personne d'une base de données n'épuise pas la possibilité qu'ont d'autres personnes, qu'ils soient concurrents ou non, de disposer et d'utiliser des mêmes données (sous réserve qu'elles puissent y accéder)⁷⁵. Par conséquent, si une entreprise, qu'il s'agisse d'un réseau d'inventaires publicitaires (ad networks), d'un distributeur ou de toute autre entreprise, collecte l'adresse personnelle, le numéro de téléphone, le sexe, la date de naissance, les revenus, les goûts et les centres d'intérêts actuels, etc., d'une personne donnée, l'utilisation de ces informations n'épuise pas les possibilités que ses concurrents puisse disposer des mêmes informations sur cette personne.

Bien qu'elle ne s'est pas référée explicitement au caractère non rival des données dans sa décision relative à l'entreprise commune Telefonica UK/Vodafone UK/Everything Everywhere, la Commission a néanmoins indiqué que « *les clients ont en général tendance à donner leurs données personnelles à de nombreux acteurs du marché qui les rassemblent et les commercialisent. Par conséquent, ce type de données est en général considéré comme un produit de base.* » (§ 543).

Le caractère non rival des données est un facteur particulièrement pertinent dans le contexte du « multi-hébergement », c'est à dire lorsque les utilisateurs ont recours à plusieurs prestataires pour fournir le même type de services, donnant ainsi accès aux données les concernant à plusieurs prestataires, aucun de ces derniers n'en ayant l'exclusivité. Le caractère non rival des données reste un facteur pertinent lorsque les consommateurs tendent vers un hébergement unique car un concurrent peut toujours avoir accès au même type d'informations que l'autre opérateur en proposant des services différents (FTC, 2014;⁷⁶ Tucker and Welford, 2014⁷⁷). Un nouvel entrant

⁷⁵ Nils-Peter Schepp and Achim Wambach, On Big Data and its Relevance for Market Power Assessment, *Journal of European Competition Law & Practice*, 2016, Vol. 7, No. 2, p. 121.

⁷⁶ « Plusieurs courtiers en données partagent les mêmes sources. Et chaque courtier en données utilise des sources multiples pour des données similaires. Par exemple, l'un des courtiers en données dans cette étude obtient des informations de contact des consommateurs de vingt sources différentes. » (p. 14).

sur le marché des réseaux sociaux peut par exemple ne pas disposer de nombreuses données relatives aux intérêts des utilisateurs potentiels par le biais de son activité de réseau social. Il peut cependant avoir accès à ces informations en interagissant avec les mêmes utilisateurs différemment, c'est-à-dire par le biais d'autres activités commerciales, comme par exemple au travers d'un site marchand permettant de collecter des données sur leurs intérêts ou sur leurs achats passés⁷⁸.

Enfin, le caractère non rival des données peut contribuer à limiter le prix des données, au moins d'un point de vue théorique, puisque les clients potentiels des courtiers en données peuvent les menacer de collecter ces données eux-mêmes. En outre, si les données ne sont pas rivales, plusieurs courtiers en données pourraient les collecter et se concurrencer pour les vendre, réduisant ainsi encore le prix de ces données tout en facilitant l'accès.

Cependant, le fait que les données soient non rivales n'implique pas qu'elles soient accessibles à tous les concurrents. En réalité, ainsi qu'il a déjà été indiqué, il existe plusieurs affaires dans lesquelles les autorités de concurrence ont considéré que, bien que les données soient non rivales, l'accès à celles-ci se faisait à un coût prohibitif, de telle sorte qu'y avoir accès était considéré comme un avantage concurrentiel significatif. Dans la partie qui suit, les difficultés pratiques pouvant limiter l'accès par un concurrent aux données sont évoquées.

bb) Facteurs pouvant limiter les possibilités d'accès aux données

La collecte directe de données peut être sujette à différents types de coûts.

⁷⁷ « Le big data n'est pas rival. En d'autres termes, la collecte d'une donnée particulière n'empêche pas d'autres sociétés de collecter des données identiques par des moyens similaires ou différents. L'emploi de prestataires multiples pour le même service (multi-hébergement de l'utilisateur) et la pratique commune des opérateurs de site web d'employer des réseaux publicitaires (ad networks) et des entreprises d'analyse multiples facilite la collecte des données pertinentes de l'utilisateur par des prestataires multiples. Par exemple, si un réseau publicitaire a déterminé que l'utilisateur d'un appareil mobile donné vit dans le Connecticut, aime voyager et possède un chien, rien n'empêche un autre réseau publicitaire d'acquérir les mêmes informations – en effet, concernant un utilisateur fréquent d'internet, il est probable que des dizaines d'entreprises créent un profil similaire. Les données redondantes sont si communes qu'elles causent des problèmes aux courtiers en données. » (pp. 3-4).

⁷⁸ Certains observateurs ont également soutenu que les informations sur les recherches passées pouvaient être récupérées non seulement par les moteurs de recherche mais également par les navigateurs et les systèmes d'exploitation (Voir Andres V. Lerner, The Role of 'Big Data' in Online Platform Competition (2014), http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2482780, p.24, citant Matt Rosoff, Yes, Bing Has Been Copying Google Search Results FOR YEARS, Business Insider, February 1, 2011, http://articles.businessinsider.com/2011-02-01/tech/29975847_1_bing-director-stefan-weitz-satyanadella-msn-toolbar).

Tout d'abord, les entreprises peuvent devoir engager des investissements significatifs pour collecter des données. L'émergence et le développement de vastes centres de données illustrent l'importance des coûts fixes qui peuvent devoir être investis pour collecter et exploiter des quantités massives de données. Le niveau de ces coûts associé à l'accumulation d'ensembles de données très importants peut empêcher les petites entreprises et les nouveaux entrants de faire usage du même volume ou de la même variété de données que les entreprises déjà établies.

Deuxièmement, les données d'utilisateurs sont souvent collectées lorsqu'ils font usage d'un produit ou d'un service. Cela est particulièrement vrai des plateformes multi-faces (telles que les moteurs de recherche ou les réseaux sociaux) qui fournissent (souvent gratuitement) des services à une première catégorie d'utilisateurs et extraient de la valeur des données collectées auprès de ces utilisateurs en fournissant d'autres produits ou services à une seconde catégorie de clients (par exemple, les espaces publicitaires vendus aux annonceurs). Par conséquent, pour accéder directement à ces types de données, un nouvel entrant doit mettre en place une plateforme capable de fournir le même type de services (ou d'autres types de services qui lui permettent de collecter des données similaires) à un nombre suffisamment élevé d'utilisateurs, ce qui peut impliquer d'importants investissements, notamment en recherche et développement. En raison des effets de réseau, d'expérience et des économies d'échelle, la constitution d'une base de clientèle suffisamment large peut ne pas être simple car la qualité du service proposé dépend de la taille de la base de clientèle (voir point 1 ci-dessus). Convaincre les utilisateurs de donner accès à leurs données personnelles peut également s'avérer difficile : en effet, certains sont réticents à transférer leurs données à caractère personnel à moins qu'ils ne bénéficient de garanties suffisantes (par exemple, un stockage sécurisé, des indications sur l'utilisation des données, la réputation du prestataire de services) ou reçoivent des services à forte valeur ajoutée. En outre, la capacité d'un nouvel entrant à se constituer une large base de clientèle peut être limitée par les coûts de transfert qui peuvent être élevés dans certains secteurs, que ce soit pour des raisons exogènes (si le service en question nécessite par exemple un apprentissage) ou en raison du comportement de l'entreprise en place, ce qui, dans ce cas, peut nécessiter un examen au regard du droit de la concurrence.

cc) Facteurs pouvant limiter l'accès aux données par l'intermédiaire de tiers

Les données ne sont pas uniquement collectées par les entreprises en tant qu'intrant pour être plus compétitives que leurs concurrents. Elles peuvent également être collectées pour être revendues sur un « marché de données ». En effet, ces vingt dernières années témoignent, en particulier aux Etats-Unis, de l'émergence d'intermédiaires⁷⁹, comme les entreprises Acxiom, Datalogix, Experian, Epsilon ou Teradata, qui peuvent collecter, stocker et analyser les données pour le compte de tiers. Ces courtiers collectent les données à partir de sources variées, notamment via leur propre technologie de collecte⁸⁰, des accords avec les éditeurs de sites web permettant à l'intermédiaire de mettre en œuvre des technologies de suivi des utilisateurs (telles que les « cookies » et « pixels »), des informations publiques (numéros de téléphone, informations disponibles sur les réseaux sociaux), des données obtenues auprès des autorités publiques et d'autres entreprises (sites web, banques, boutiques en ligne, autres courtiers en données). Les clients de ces courtiers en données sont des entreprises actives dans différents secteurs et peuvent inclure d'autres opérateurs collectant également de larges ensembles de données. Certains courtiers en données ont par exemple conclu des partenariats avec des sociétés comme Facebook ou Google afin de les aider à améliorer le ciblage de leurs offres publicitaires⁸¹. Réciproquement, un opérateur actif sur un marché reposant en partie sur l'exploitation des données peut juger rentable de donner à d'autres opérateurs accès aux données qu'il a collectées, afin d'en extraire une source de revenus supplémentaires.

L'utilisation des données de tiers peut constituer une alternative à la collecte directe de données. En effet, cet accès par le biais d'intermédiaires peut être moins coûteux : les coûts fixes de la collecte de données sont mutualisés sur un plus grand

⁷⁹ Les courtiers en données sont apparus il y a déjà quelques temps. Acxiom, par exemple, société qui vend des données physiques et numériques, a été créée en 1969.

⁸⁰ Par exemple, Twitter a créé une plateforme donnant accès au flux de tweets et aux données brutes associées. Dans ce contexte, Twitter a d'abord conclu des partenariats avec les revendeurs de données (GNIP, Datasift et NTT), qui ont agi comme intermédiaires avec les sociétés utilisant les tweets. Après l'acquisition de GNIP en 2014, Twitter a décidé de mettre fin à ces partenariats pour établir une relation commerciale directe avec les sociétés acquérant ces données. Facebook a également conclu un partenariat en mars 2015 avec Datasift concernant le traitement de données thématiques provenant de ses utilisateurs, qui seront mises à disposition des annonceurs par le biais d'une plateforme Datasift.

⁸¹ Facebook utilise par exemple les données de courtiers en données (y compris les pages web visitées, les abonnements aux newsletters, l'argent dépensé en ligne et hors ligne, etc.) pour enrichir ses propres ensembles de données sur les amis et les « likes » de ses utilisateurs (Evelyn M. Rusli, Buy Signal: Facebook Widens Data Targeting, <http://www.wsj.com/articles/SB10001424127887324504704578412960951909032>).

nombre d'entreprises utilisatrices et une entreprise cliente peut n'acquérir que les données dont elle a besoin en termes de volumes et de variété sans avoir à supporter d'importants coûts fixes. En outre, les services proposés par les intermédiaires sont nombreux et peuvent inclure des prestations d'analyse de données, ce qui réduit encore les coûts fixes associés à l'exploitation des données. Recourir à un intermédiaire peut également aider une entreprise à étendre les volumes et/ou la variété de ses propres bases de données ou la qualité de ses services d'exploitation de données.

Les données de tiers présentent cependant certains inconvénients. Tout d'abord l'étendue des données accessibles par l'intermédiaire des courtiers en données peut être limitée relativement à ce que peuvent obtenir les principaux opérateurs sur un marché lié aux données. Ainsi, les données comportementales collectées par le biais de cookies ou de pixels peuvent être d'une portée limitée et ainsi s'avérer être des substituts imparfaits aux données individuelles qui sont directement renseignées par l'utilisateur d'un service, car les cookies ne renseignent que les sites visités par une personne mais ne permettent pas d'obtenir des informations sociodémographiques détaillées pourtant susceptibles de présenter une grande valeur. Facebook élabore ainsi des profils détaillés de ses utilisateurs grâce aux informations personnelles très sensibles touchant à tous les domaines de la vie (famille, enseignement, emploi, intérêts politiques, loisirs, etc.) qu'ils renseignent lorsqu'ils utilisent le réseau social. En outre, l'exploitation des données peut offrir des « rendements de gamme » croissants, la diversification des données et leur combinaison apportant de nouvelles perspectives, par exemple en termes de profilage du client. Il semblerait ainsi que Google bénéficie non seulement de l'importante quantité de données collectées par l'intermédiaire des requêtes sur ses moteurs de recherche mais également des données collectées sur ses autres services tels que la messagerie électronique, les services vidéo, les téléphones, etc. Par conséquent, si les courtiers en données sont contraints dans les types de données qu'ils peuvent collecter, ceci peut conduire à diminuer la valeur relative de leurs données par rapport aux données propres de l'entreprise.

Par ailleurs, certains coûts spécifiques engagés par les intermédiaires en données peuvent être répercutés sur leurs clients⁸². Ces contraintes techniques peuvent être

⁸² L'OCDE souligne par exemple que « la fourniture de données de qualité supérieure peut requérir des investissements immédiats et à long terme significatifs avant que les données ne puissent être

d'autant plus difficiles à surmonter lorsque la valeur des données tend à diminuer rapidement avec le temps, comme c'est notamment le cas des données de localisation, la solution technique devant permettre une mise à jour fréquente de la base fournie à l'acheteur des données. Ces différents facteurs expliquent pourquoi, ainsi que l'OCDE l'expose, « *les créateurs et contrôleurs de données ne sont pas nécessairement poussés à partager leurs données. L'une des raisons en est que les coûts du partage des données sont perçus comme supérieurs aux avantages attendus du partage* »⁸³. En outre, l'intégration des utilisateurs de données avec les collecteurs de données peut générer différentes efficiences relativement au recours à des courtiers en données indépendants. Ainsi, dans sa décision concernant la fusion TomTom/Tele-Atlas, la Commission a souligné que l'intégration entre TomTom (fabricant d'appareils portables de navigation) et Tele-Atlas (vendeur de bases de données cartographiques) pouvait permettre à Tele-Atlas de bénéficier des retours d'expérience des utilisateurs (telles que les corrections des erreurs) rassemblés par TomTom par le biais de sa large base de clientèle (§ 246), bien que l'ampleur de ces efficiences restait incertaine.

Troisièmement, les entités qui collectent des données de valeur peuvent être contraintes, en vertu de dispositions contractuelles ou légales, dans les modalités de partage de ces données avec des tiers. En effet, en collectant des données à caractère personnel, l'opérateur garantit en général aux utilisateurs que leurs données personnelles ne seront pas communiquées à un tiers sans leur accord. Sans une telle garantie, les utilisateurs peuvent être réticents à communiquer leurs données personnelles. En tout état de cause, les règles de protection de la confidentialité dans la plupart des pays limitent strictement la fourniture de données personnelles à des tiers à des fins commerciales. En réalité, plus les données sont sensibles, moins elles sont susceptibles d'être collectées ou transmises à des intermédiaires spécialisés dans les données qui, pour la plupart, sont peu connus des consommateurs et ont moins d'incitation à défendre leur réputation en matière de protection de la vie privée, et auxquels les consommateurs font donc moins

partagées », y compris « i) la mise en données, ii) la collecte des données, iii) le « nettoyage » des données et iv) la préservation des données (...), les modèles de données et les algorithmes de stockage et de traitement des données, et même les infrastructures informatiques sécurisées pour le stockage, le traitement et l'accès (partagés) aux données ».

⁸³ OECD, *Data-Driven Innovation: Big Data for Growth and Well-Being*, OECD Publishing, Paris (2015), <http://dx.doi.org/10.1787/9789264229358-en>, p.192.

confiance⁸⁴. Dans l'ensemble, le rôle joué par les courtiers en données peut varier en fonction du pays, du type de données concernées et du niveau de protection de la confidentialité des données personnelles dont bénéficient les utilisateurs en vertu des dispositions nationales.

Enfin, les incitations qu'ont les opérateurs actifs sur le marché à partager leurs données avec leurs concurrents pourraient s'avérer insuffisantes. Si les données ont une valeur, leur communication aux concurrents pourrait en effet réduire significativement l'avantage concurrentiel de l'entreprise qui les collecte dans des proportions telles que ses revenus, *in fine*, n'augmentent pas. Par conséquent, la plupart des entreprises qui sont actives sur les marchés sur lesquels elles collectent des données sont souvent réticentes à les communiquer à leurs concurrents. Dans certains cas, cette réticence peut aller jusqu'à imposer des restrictions d'accès aux données pour les tiers. Par exemple, Graef et al. (2015) rapporte que Facebook empêche d'autres entreprises de collecter des données sur les utilisateurs de son réseau social sans son consentement⁸⁵. Sur un plan plus général, il existe plusieurs marchés dans lesquels les concurrents se plaignent d'un manque d'accès aux données pertinentes, indépendamment du fait que cette pratique puisse ou non être considérée comme anticoncurrentielle⁸⁶.

En conclusion, le fait que les données aient un caractère non rival ne résout pas l'ensemble des problèmes de concurrence susceptibles d'être associés aux données. En effet, la non rivalité n'implique pas nécessairement que les données soient accessibles à tous les concurrents ou, plus précisément, que tous les concurrents

⁸⁴ Voir également OECD, *Data-Driven Innovation: Big Data for Growth and Well-Being*, OECD Publishing, Paris (2015), <http://dx.doi.org/10.1787/9789264229358-en>, qui fait valoir que « *de meilleurs régimes de gouvernance des données sont nécessaires pour surmonter les obstacles à l'accès aux données, à leur communication et à leur interopérabilité. Ces régimes peuvent avoir un impact sur les incitations à partager les données et la possibilité de les utiliser de manière interopérable. Les éléments à prendre en compte pour une gouvernance efficace des données incluent l'accès aux données et leur réutilisation ; la portabilité et l'interopérabilité ; le lien et l'intégration ; la qualité et la préservation ; la « propriété » et le contrôle ; et la valeur et la tarification.* »

⁸⁵ Inge Graef/Sih Yuliana Wahyuningtyas/Peggy Valcke, *Assessing data access issues in online platforms* (2015), http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2647309, se réfère à la Déclaration des droits et des responsabilités de Facebook qui prohibe la collecte automatique des contenus de l'utilisateur : « Vous n'obtiendrez pas d'informations concernant des internautes ni les contenus qu'ils publient, et n'accéderez pas à Facebook à l'aide de méthodes automatisées (telles que robots, araignées, etc.) sans notre permission préalable » disponible sur <https://www.facebook.com/legal/terms>.

⁸⁶ Cf. les marchés de l'énergie mentionnés plus haut. D'autres cas concernent le transfert des informations techniques aux opérateurs des marchés d'après-vente (marchés de maintenance des produits, de formation des salariés, etc.).

soient également en mesure de collecter ces données. Une analyse au cas par cas s'avère dès lors nécessaire.

b) La disponibilité des données numériques

aa) Présentation

Les volumes de données qui peuvent être collectées n'ont jamais été aussi importants qu'actuellement. Ainsi que l'explique l'OCDE, le passage de nombreuses activités à l'ère numérique et le déploiement de services connectés pour des activités hors ligne entraînent une augmentation massive de la quantité de données générées par les utilisateurs⁸⁷ et susceptibles d'être collectées par les entreprises. Dans ce contexte, il est dit que les données « sont partout », ce qui réduit ainsi le risque qu'elles puissent conférer du pouvoir de marché aux entreprises.

L'argument tiré de « l'omniprésence des données » est d'autant plus pertinent que la valeur des données dépend généralement des informations qui peuvent en être extraites, et non des données elles-mêmes. Si le même type d'informations peut être extrait à partir de différents types de données qui peuvent être obtenues par des biais différents, le risque qu'une entreprise ne soit pas en mesure d'accéder aux mêmes informations que ses concurrents est moindre.

Graef (2015) donne ainsi l'exemple d'un moteur de recherche qui connaît les préférences musicales d'un utilisateur grâce aux requêtes qu'il a faites et d'un réseau social obtenant le même niveau de connaissance en consultant les informations de profil que l'utilisateur a partagées sur sa plateforme⁸⁸. Un rapport remis au Président des Etats-Unis⁸⁹ indique à cet égard que l'analyse des réseaux sociaux peut permettre un ciblage publicitaire potentiellement aussi efficace que celui réalisé à partir des données de recherche. Lerner (2014) avance également que les données collectées par Amazon sur les achats effectués par les consommateurs pourraient être aussi efficaces que celles détenues par Google pour améliorer le ciblage publicitaire. De même, les fournisseurs d'accès à Internet sur mobile pourraient accéder à des volumes de données considérables sur la localisation en temps réel

⁸⁷ OECD, Data-Driven Innovation: Big Data for Growth and Well-Being, OECD Publishing, Paris (2015), <http://dx.doi.org/10.1787/9789264229358-en>

⁸⁸ Inge Graef, Market definition and market power in data: the case of online platforms, *World Competition: Law and Economics Review*, Vol. 38, No. 4 (2015)

⁸⁹ Executive Office of the President, President's Council of Advisors on Science and Technology, , Report to the President - Big Data and Privacy: a Technological Perspective (2014), https://www.whitehouse.gov/sites/default/files/microsites/ostp/PCAST/pcast_big_data_and_privacy_-_may_2014.pdf

des consommateurs, proches de celles que Google et Facebook peuvent détenir⁹⁰. En outre, certaines entreprises, notamment lorsqu'elles ne sont pas concurrentes, peuvent accepter de procéder à un échange croisé des données qu'elles collectent afin d'augmenter leur niveau d'information.

Les autorités de concurrence américaines et européennes ont considéré, dans certains cas de concentration, qu'un meilleur accès à certaines données grâce à l'opération ne posait pas de problèmes de concurrence en raison des quantités importantes de données qui restaient disponibles pour les concurrents. Ainsi, dans sa décision Google/DoubleClick de 2008, la Commission européenne a-t-elle considéré que « *la combinaison des données sur les recherches et des données sur le comportement de navigation sur le web des utilisateurs [générée à la suite de la fusion] est déjà accessible à un certain nombre de concurrents de Google aujourd'hui* ». La décision précise en outre que « *les concurrents peuvent également acquérir les données ou les services de ciblage auprès de tiers* » et que « *les données sont également disponibles auprès des fournisseurs d'accès à Internet* » (§ 365). Un raisonnement similaire a été tenu par la FTC : « *un certain nombre de concurrents de Google disposent de masses de données de valeur auxquelles Google n'a pas accès. Par exemple, les principaux concurrents de Google sur le marché de l'intermédiation publicitaire, Microsoft, Yahoo!, et Time Warner ont accès à leurs propres répertoires uniques de données. Ces entreprises détiennent des moteurs de recherche appréciés et ont accès à des informations sur les consommateurs à partir de leurs serveurs publicitaires internes, de leurs services d'intermédiation publicitaire, des autres sites qu'elles détiennent et des logiciels. L'entrée et l'expansion de ces concurrents bien financés a transformé le marché de l'intermédiation publicitaire ces six derniers mois. Toutes ces entreprises sont verticalement intégrées, et toutes semblent bien positionnées pour concurrencer vigoureusement Google sur ce nouveau marché.* » (pp. 12-13).

Bien qu'on puisse considérer que ces décisions sont trop anciennes pour être toujours pertinentes au vu de l'évolution rapide qui caractérise ce secteur, des décisions plus récentes présentent le même raisonnement. Dans sa décision sur l'entreprise commune Telefonica UK/Vodafone UK/ Everything Everywhere de 2012,

⁹⁰ Geoffrey A. Manne and Ben Sperry, The problems and perils of bootstrapping privacy and data into an antitrust framework, CPI Antitrust Chronicle, May 2015, p.11, également sur http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2617685

la Commission a estimé que l'entreprise commune n'aurait pas accès à une base de données pour la publicité mobile qu'aucun concurrent ne pourrait égaler puisque « *les informations accessibles par la [nouvelle entité] sont également accessibles dans une large mesure par les acteurs du marché, existants comme nouveaux, tels que Google, Apple, Facebook, les émetteurs de cartes, les agences de référence ou les distributeurs. (...) Les clients ont en général tendance à communiquer leurs données à caractère personnel à de nombreux acteurs du marché qui les rassemblent et les commercialisent. Par conséquent, ce type de données est en général considéré comme un produit de base. (...) Il existe d'autres manières d'atteindre un grand nombre de consommateurs, tels que les fournisseurs de services de base par exemple [eau, énergie, etc.]. En outre, les fournisseurs d'accès à Internet semblent également être en position d'atteindre un très grand nombre de consommateurs* » (§§ 543-544). En 2014, les autorités de concurrence américaines et européennes ont appliqué un raisonnement similaire pour autoriser sans condition l'acquisition de WhatsApp par Facebook. La Commission européenne a notamment expliqué qu' « *il existe un nombre significatif d'acteurs sur le marché qui, à côté de Facebook, collectent les données d'utilisateurs. Ces acteurs incluent tout d'abord Google, qui représente une part importante des données d'utilisateur sur Internet, ainsi qu'Apple, Amazon, eBay, Microsoft, AOL, Yahoo, Twitter, IAC, LinkedIn, Adobe et Yelp entre autres.* » (§ 188). De plus, « *il continuera à y avoir une quantité importante de données d'utilisateur sur Internet qui présentent un intérêt d'un point de vue publicitaire, et qui ne sont pas sous le contrôle exclusif de Facebook.* » (§ 189).

bb) Discussion : de la disponibilité des données à la possibilité de substitution entre les données

Les volumes et la variété des données augmentent de manière constante depuis des années. Cela vaut tant pour l'environnement en ligne que pour le monde physique grâce aux objets connectés. Cela étant dit, la pertinence de l'argument selon lequel les données « *sont partout* » dépend de manière décisive de l'accessibilité des données (voir ci-dessus) et du caractère substituable des différentes catégories de données. Cette substituabilité peut toutefois s'avérer difficile à mesurer.

Par exemple, comme indiqué plus haut, les informations obtenues par les réseaux sur leurs utilisateurs peuvent être aussi riches, ou même plus riches que celles obtenues par les moteurs de recherche via les requêtes de leurs utilisateurs.

Cependant, contrairement aux données produites sur les réseaux sociaux, les données collectées par les moteurs de recherche ont l'avantage de permettre d'identifier les consommateurs qui recherchent activement un service ou un bien donné. Une incertitude similaire pourrait prévaloir en comparant les données collectées par les distributeurs en dur (obtenues par les cartes de fidélité par exemple ou par l'extraction des données des factures) et celles obtenues par les distributeurs en ligne, ces dernières incluant les produits qu'un utilisateur a envisagé d'acheter avant de procéder à un achat effectif. De même, les données collectées et utilisées par le biais des PC et des ordinateurs portables ne seraient pas nécessairement substituables aux données collectées et utilisées sur les téléphones portables puisque seules les données contenant la localisation de l'utilisateur en temps réel pourraient être adaptées aux téléphones portables⁹¹.

La substituabilité de différentes catégories de données a été examinée dans plusieurs décisions passées, sans lien cependant avec les marchés dits numériques. C'est ainsi que dans sa décision dans l'affaire TomTom/Tele Atlas, déjà mentionnée (décision M.4854, § 22), la Commission a conclu que les bases de données cartographiques numériques ne servant pas à la navigation n'étaient pas substituables aux bases de données destinées à la navigation car ces dernières doivent inclure davantage d'informations (tels que le type de route, les informations sur la circulation, les restrictions sur les virages).

En évaluant le caractère substituable des données, les autorités de concurrence prennent également en considération les possibilités de substitution du côté de l'offre et les barrières à l'entrée. Une affaire dans laquelle ces critères ont été particulièrement étudiés est la décision relative à l'opération Thomson Corporation/Reuters Group (décision M.4726). Cette opération donnait lieu à des chevauchements significatifs entre certaines bases de données et ensembles de contenus sur plusieurs marchés de services financiers. Les préoccupations de concurrence étaient liées aux parts de marché importantes de la nouvelle entité ainsi qu'aux coûts et au temps nécessaire pour créer des bases de données similaires à celles proposées par l'entité résultant de l'opération. En effet, ces données

⁹¹ Certains analystes estiment dans ce contexte que les fournisseurs d'accès à Internet sur mobiles ont accès à une quantité de données bien plus importantes sur la localisation en temps réel des consommateurs, proche des données dont Google et Facebook peuvent bénéficier. Voir Geoffrey A. Manne and Ben Sperry, *The problems and perils of bootstrapping privacy and data into an antitrust framework*, CPI Antitrust Chronicle, May 2015, p.11, également sur http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2617685.

regroupent des contributions de nombreux courtiers et analystes financiers différents - qui peuvent avoir des accords exclusifs avec Thomson ou Reuters ou qui peuvent n'accepter de partager leurs recherches avec un distributeur que si ce dernier dispose d'une base de clientèle significative - contributions qui sont récupérées par le biais d'une infrastructure logicielle spécifique et peuvent donc devoir être standardisées pour présenter un attrait pour les clients.

Dans la plupart des affaires de concentration examinées par les autorités de concurrence en lien avec des marchés reposant sur la collecte et l'exploitation de données, les parties concernées par l'opération n'utilisaient leurs données qu'en interne, sans les vendre à des tiers. Par conséquent, elles n'étaient pas présentes sur un éventuel marché de données. Les contours d'un éventuel marché de données n'ont dès lors pas dû être définis.

Cependant, ainsi qu'il est expliqué plus haut, dans ses décisions Google/DoubleClick, Facebook/Whatsapp, *Telefónica/Vodafone/Everything Everywhere*, ainsi que celle relative à la fusion *Publicis/Omnicom*, la Commission a eu tendance à considérer qu'en dépit des moyens étendus de collecte des données induits par ces opérations, la quantité de données accessibles et utilisables à des fins publicitaires ou aux fins d'analyses resterait suffisante pour que les avantages que tiraient les parties concernées des opérations en cause ne soient pas irrattrapables. Inversement, dans la décision TomTom/Tele Atlas, la Commission a estimé que la production d'une base de données cartographiques à des fins de navigation serait très coûteuse car une des caractéristiques de ce type de base de données est qu'elle est compilée manuellement par une flotte de véhicules et mise à jour régulièrement (§ 24). La Commission est arrivée à la même conclusion dans sa décision Thomson/Reuters. Ensemble, ces considérations pourraient suggérer que les données seraient effectivement plus faciles à collecter sur les marchés numériques, ce qui limiterait les préoccupations de concurrence. Plus généralement, il faut toutefois garder à l'esprit que le Big data n'en est encore qu'à ses débuts et que des développements à venir dans les outils et modalités d'exploitation des données sont probables. Les données qui pourraient apparaître aujourd'hui comme des substituts ne le seront peut-être plus à l'avenir, et réciproquement. Les distinctions possibles entre les données hors ligne et en ligne (voir ci-dessus), entre les données générées par le biais des applications mobiles et celles collectées par des appareils fixes, entre les données récupérées auprès des réseaux sociaux et

celles provenant des requêtes de recherche pourraient devenir plus évidentes à mesure que les moyens d'exploiter ces différents types de données se développent. En revanche, certaines données telles que les informations de cartographie pour la navigation dont la collecte est coûteuse, car elles ne peuvent pas être téléchargées automatiquement via Internet ou autrement, se verront potentiellement à l'avenir (par exemple, pour des raisons tenant à la qualité des données), remplacer par des données activement et volontairement renseignées par un grand nombre de personnes, ou même par leurs voitures. S'agissant des données cartographiques, une telle ressource existe sous la forme des données collectées pour le projet Open Street Map⁹².

Certaines des décisions susmentionnées laissent déjà la place à une analyse plus approfondie de la substituabilité entre données. Ainsi, la Commission européenne a déjà procédé à une distinction claire entre les services publicitaires hors ligne et en ligne dans ses décisions Google/DoubleClick et Microsoft/Yahoo, notamment en raison des différentes possibilités de ciblage offertes par la publicité en ligne (voir notamment Google/DoubleClick, §§ 50-52 ; Microsoft/Yahoo!, §§ 64-65). Mais, plus récemment, dans sa décision *Telefónica/Vodafone/Everything Everywhere*, la Commission a noté qu'une vaste majorité des personnes interrogées à l'occasion du test de marché considéraient que « *la fourniture de services d'analyses de données pour la publicité en ligne statique n'est pas substituable avec la fourniture de services d'analyses de données pour la publicité mobile. (...) Les deux services collectent en effet un type différent d'informations et une quantité différente de détails relatifs au consommateur. Les informations générées par l'analyse des données mobiles sont en général plus personnelles, géo-localisées et peuvent faire l'objet d'un croisement avec les comportements d'appels, ce qui ne peut pas être proposé, dans une mesure comparable, par l'analyse des données en ligne. Les services d'analyses de données en ligne et mobiles sont en général considérés comme complémentaires et continueront à être utilisés parallèlement* » (§ 200). La Commission a cependant considéré que les données relatives à la géolocalisation des clients sont suffisamment répandues pour que la nouvelle entité ne dispose pas d'avantage concurrentiel sur ses concurrents : « *parmi les acteurs les plus connus qui rassemblent ce type des informations, figurent Apple, Facebook ou Google (avec Google Maps)* » (§ 545) ainsi que Microsoft.

⁹² www.openstreetmap.org .

De même, dans sa décision Facebook/Whatsapp, la Commission a observé « *qu'un certain nombre de personnes interrogées considèrerait que d'autres formes de publicité non liée aux recherches en ligne ne sont pas aussi efficaces que la publicité sur les sites de réseaux sociaux et notamment sur Facebook, en raison de sa large base d'utilisateurs, très investis dans l'utilisation du réseau, et de ses opportunités de ciblage des annonces publicitaires* » (§ 77, soulignement ajouté), ce qui indique possiblement que les données collectées par Facebook permettent un meilleur ciblage publicitaire que celles collectées par d'autres opérateurs sur le marché des services publicitaires en ligne.

En conclusion, le fait que les coûts de collecte des données aient fortement diminué en raison du développement des activités en ligne, mobiles, et hors ligne connectées n'implique pas que tous les types de données soient interchangeables.

4. Aspects liés au volume et à la variété des données collectées

a) Présentation

L'importance de l'avantage concurrentiel découlant de la collecte et de l'exploitation des données est susceptible de dépendre, de manière décisive, de la nécessité ou non de disposer d'un large volume et/ou d'une large variété de données, ce qui ne serait à la portée que d'une entreprise établie, de grande taille et/ou diversifiée. Si cette nécessité n'est pas vérifiée, il pourrait alors être affirmé que les concurrents peuvent facilement obtenir le volume de données nécessaires pour être sur un pied d'égalité, notamment au vu de la disponibilité des données sur les marchés numériques ou connectés et de leur caractère non rival.

A cet égard, Lerner (2014) avance que la valeur marginale des données utilisées à des fins d'inférence peut diminuer rapidement lorsqu'un certain volume de données a déjà été collecté. Les données pourraient ainsi afficher des rendements marginaux décroissants en fonction du volume de données collectées, ce qui limiterait les avantages concurrentiels résultant de grandes quantités de données⁹³.

Cette affirmation pourrait être particulièrement pertinente pour certaines utilisations spécifiques de données telles que l'alimentation d'un algorithme de moteur de recherche ou, plus généralement, lorsque les données sont utilisées pour faire des prédictions. Les prédictions sont en effet basées sur des probabilités. Plus les

⁹³ Nils-Peter Schepp and Achim Wambach, On Big Data and its Relevance for Market Power Assessment, *Journal of European Competition Law & Practice*, 2016, Vol. 7, No. 2, p. 121.

observations sont nombreuses, plus la mesure d'une probabilité et la prédiction seront fiables. L'augmentation de la précision de la prédiction décroît cependant lorsque la taille de la base de données augmente : le biais d'échantillonnage statistique associé à toute extrapolation à partir d'un ensemble de données décroît lorsque le nombre d'observations augmente, mais à un rythme de plus en plus faible à mesure que la taille de l'échantillon augmente⁹⁴.

Par exemple, la plupart des moteurs de recherche prennent en compte les choix observés de leurs utilisateurs parmi les résultats présentés en réponse à une requête, afin d'améliorer et de rendre plus pertinent le classement de ces résultats à l'avenir, pour des requêtes similaires. Si le nombre des observations est relativement élevé, le moteur de recherche aura un échantillon d'observations plus représentatif auquel se référer afin d'améliorer le classement des résultats, avec un risque limité d'erreurs liées à des observations inhabituelles. Cependant, à mesure que le nombre de requêtes de recherche augmente, la valeur marginale des informations de chaque requête de recherche diminue.

Le nombre de requêtes nécessaires à un moteur de recherche pour être compétitif peut ainsi être inférieur au nombre cumulé de recherches réalisées sur le moteur de Google. Certaines considérations de la décision Microsoft/Yahoo tendent à soutenir cette position, la Commission indiquant que « *Microsoft a présenté une étude de référencement qui compare la pertinence algorithmique de ses résultats de recherches à celle de Google et Yahoo [...]. Les résultats de cette étude illustrent que Microsoft et Yahoo [...]. Pour les requêtes les plus fréquentes cependant, l'écart global de pertinence entre les moteurs est très faible [...]* » (§ 166).

En outre, il faut garder à l'esprit que la capacité à extraire les informations à partir des données ne repose pas exclusivement sur les quantités de données disponibles, mais également sur les algorithmes qui analysent les données, qui ne sont pas tous de même qualité⁹⁵. Les différents niveaux de compétitivité ou de qualité des offres sur un marché reposant sur l'exploitation des données ne sont pas totalement imputables à la taille plus ou moins grande des bases de données constituées par les concurrents sur le marché. Afin d'illustrer ce point, la Commission a conclu dans sa décision Microsoft/Yahoo que « *bien que (...) Google semble montrer une*

⁹⁴ Par exemple, voir Andres V. Lerner, The Role of 'Big Data' in Online Platform Competition (2014), http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2482780, p. 35.

⁹⁵ Nils-Peter Schepp and Achim Wambach, On Big Data and its Relevance for Market Power Assessment, *Journal of European Competition Law & Practice*, 2016, Vol. 7, No. 2, p. 122.

meilleure performance en termes de pertinence notamment pour [...] les requêtes, cela ne prouve pas que la pertinence pour l'utilisateur est fonction de l'échelle des données, puisque les études ci-dessus ne prennent pas en compte la technologie des différents moteurs de recherche qui ne dépend pas des volumes de données » (§ 168). Le rôle des analyses de données dans la compensation (ou l'exacerbation) de la petite taille d'une base de données ne doit donc pas être sous-estimé.

Un autre facteur limitant l'avantage que retirent les plus grandes entreprises de la détention de bases de données volumineuses tient à la valeur parfois fortement décroissante des données dans le temps. Les données historiques par exemple, bien qu'elles soient utilisées pour l'analyse des tendances dans les marchés publicitaires, peuvent avoir une valeur relativement faible pour une prise de décision instantanée telle que le choix de l'annonce à présenter dans les enchères en temps réel (*real-time bidding*). En outre, les données historiques peuvent avoir une valeur relativement faible pour certains acteurs tels que les moteurs de recherche au vu du taux élevé de nouvelles requêtes: ainsi que le rapporte Google, 15 % des requêtes présentées chaque jour sont nouvelles (voir Lerner (2014), § 64), ce qui implique que les algorithmes ont en permanence besoin de nouvelles données pour être efficaces en fournissant, pour ces nouvelles requêtes, le classement de résultats le plus pertinent⁹⁶.

b) Discussion

A titre préliminaire, des commentateurs⁹⁷ considèrent que les stratégies suivies par certains acteurs majeurs concernant leurs choix d'investissement et leurs pratiques envers les données d'utilisateurs démontrent qu'ils souhaitent accéder à une quantité de données encore plus importante que celle dont ils disposent actuellement, ce qui suggère que le volume et la variété des données collectées sont des paramètres clés de la concurrence. Cependant, il peut y avoir plusieurs motifs à l'acquisition d'une société (tels que l'acquisition d'une large base d'utilisateurs permettant de générer des revenus publicitaires ou auprès desquels valoriser de nouveaux services). Par ailleurs, le souhait apparent de certaines sociétés de collecter, combiner, stocker et utiliser des volumes toujours plus importants de données pourrait simplement

⁹⁶ A noter cependant que Google peut rester avantagé par rapport aux nouveaux entrants car il dispose d'un pool d'utilisateurs plus vaste procédant à de nouvelles recherches et pourra ainsi améliorer son moteur de recherche plus rapidement que ses concurrents.

⁹⁷ Voir Nathan Newman, Search, Antitrust and the Economics of the Control of User Data, Yale Journal on Regulation, Vol. 30, No. 3, 2014, also available at SSRN: <http://ssrn.com/abstract=2309547> or <http://dx.doi.org/10.2139/ssrn.2309547>.

indiquer que les coûts marginaux de collecte des données sont très faibles. Ainsi, à moins que certains coûts fixes importants ne soient engagés ou que des risques juridiques soient encourus en raison du traitement de ces données supplémentaires, les volumes toujours plus importants de données collectées ne constituent pas en eux-mêmes la preuve qu'une collecte de données de vaste ampleur soit indispensable pour être compétitif sur les marchés concernés.

Cependant, bien que plusieurs des arguments développés ci-dessus fassent ressortir les limites, dans certaines circonstances, des avantages que retire une entreprise du volume et de la variété des données qu'elle détient, leur validité ne doit pas être admise pour tous les cas.

S'agissant tout d'abord de l'impact de l'obsolescence des données, il importe de garder à l'esprit que, bien que la valeur de certaines données puisse diminuer relativement rapidement avec le temps, toutes les données n'ont pas une valeur exclusivement temporaire. En fonction du marché, certaines données, telles que le sexe, le nom, l'adresse, la date de naissance, l'emploi, etc., peuvent ne pas perdre de valeur avec le temps. Ainsi, une société disposant de ces données peut avoir un avantage durable sur ses concurrents. En outre, si les données collectées doivent être mises à jour très fréquemment, les entreprises sont alors susceptibles de devoir collecter des volumes de données dans un délai très court, ce qui augmente ainsi la contrainte représentée par les données.

Deuxièmement, la valeur marginale décroissante des données mise en avant par Lerner et d'autres dans le cas des moteurs de recherche et des systèmes de recommandation découle du fait que ces données sont utilisées à des fins d'inférence. Toutes les données ne sont cependant pas utilisées uniquement à des fins d'inférence. Dans les marchés plus « classiques » par exemple, les bases de données peuvent contenir des coordonnées ou des informations de consommation. Dans ces cas, chaque observation intégrée à la base de données peut revêtir la même importance, car chacune représente un nouveau prospect.

Troisièmement, les entreprises, afin de procéder à des inférences informatives, s'intéressent tant à la qualité des inférences qu'elles peuvent faire en s'appuyant sur une base de données spécifique qu'à la quantité des observations contenues dans cette base. Une base de données de petite taille peut ainsi avoir une dimension suffisante pour permettre à un annonceur ou à une agence publicitaire de faire des

inférences sur les probabilités d'achat par les consommateurs d'un bien donné. Cependant, pour que cette inférence ait de la valeur, l'entreprise a besoin de disposer d'informations correspondantes sur un large ensemble de personnes. Par exemple, grâce à une base de données relativement limitée, une entreprise peut avoir inféré que les femmes brunes achètent un certain type de maquillage. Pour que cette inférence soit utile, l'entreprise a cependant besoin d'une grande base de données décrivant la couleur des cheveux de ses clients potentiels.⁹⁸

Quatrièmement, l'existence de rendements d'échelle décroissants, s'agissant des données, ne préjuge pas de l'existence d'un seuil de données en deçà duquel une entreprise peut être désavantagée par rapport aux entreprises établies. Ce niveau peut être élevé dans certaines activités. En examinant par exemple la publicité liée aux recherches par opposition au moteur de recherche lui-même, la Commission a observé, dans sa décision Microsoft/Yahoo (2010) que presque tous les annonceurs estimaient qu'une base importante d'utilisateurs était nécessaire pour exercer une concurrence effective dans le secteur de la publicité liée aux recherches puisqu'elle assure une exposition plus large⁹⁹.

Cinquièmement, le volume de données nécessaires pour faire une inférence peut être relativement faible. Cependant, sur certains marchés, le nombre d'inférences qui doivent être faites tous les jours est considérable, ce qui augmente fortement le volume de données nécessaire. Par exemple, les requêtes adressées à un moteur de recherche sont très diverses et le nombre de requêtes différentes très élevé. Par conséquent, une proportion significative de requêtes peut ne se rencontrer que rarement au niveau d'un seul moteur de recherche. Dans ce contexte, le faible nombre d'utilisateurs de petits moteurs de recherche procédant à des requêtes rares ne leur permet pas d'améliorer la pertinence de leur classement, ce qui pousse les consommateurs à opter pour un autre moteur de recherche, plus utilisé, pour l'ensemble de leurs requêtes. Le fait que les requêtes et les comportements des utilisateurs évoluent à un rythme rapide rend d'autant plus décisive le nombre de

⁹⁸ Voir D. Evans, *The online advertising industry: economics, evolution and privacy*, *Journal of Economic Perspectives*, 2009, 23 (3), pp. 37-60.

⁹⁹ « L'enquête de phase I tend à soutenir le point de vue de la partie notificante que l'échelle importe pour être un concurrent efficace. Presque tous les annonceurs répondant à l'enquête de marché considèrent que Microsoft ne dispose pas d'un volume de trafic suffisant pour constituer une alternative attractive à Google » (§153). « Tous les annonceurs répondant à l'enquête de marché considèrent que Google est « incontournable » pour les campagnes publicitaires de « search ». De plus, tous les annonceurs répondant à l'enquête de marché soulignent le fait que le volume de requêtes de Google est l'une des principales raisons pour lesquelles Google est un « incontournable » » (§157).

requêtes reçues par un moteur de recherche, afin d'identifier ces évolutions, d'adapter les résultats de recherche et d'améliorer ainsi leur pertinence dans un délai réduit¹⁰⁰.

La même sorte de raisonnement pourrait s'appliquer à d'autres catégories de sites (par exemple les recommandations d'achats en ligne). Par conséquent, une échelle relativement importante de collecte de données peut contribuer à la compétitivité d'une plateforme en ligne.

Enfin, en fonction de l'utilisation des données, la variété des données contenues peut être aussi importante que l'échelle d'une base de données : la valeur d'une base de données ne repose pas uniquement sur le nombre de personnes qu'elle décrit mais également sur la quantité d'informations collectées sur chaque personne. A ce titre, la collecte de données par l'offre de différents services permet à l'entité qui procède à la collecte de rassembler des connaissances sur de multiples aspects du comportement et des préférences de ses utilisateurs. Tandis que chacun de ces aspects peut avoir un intérêt en soi pour un annonceur ou un distributeur, leur combinaison peut conduire à une meilleure connaissance de chaque utilisateur afin, par exemple, de mieux inférer la probabilité qu'un utilisateur soit susceptible d'être intéressé par l'achat d'un certain produit à un moment donné.

La valeur retirée de la variété des données a été soulignée dans plusieurs rapports. Le rapport remis au Président des Etats-Unis indique ainsi que « [l'association de données provenant de différentes sources] *peut dévoiler de nouvelles significations. Notamment, la fusion des données peut se traduire par l'identification des personnes, la création de profils d'individus et le suivi des activités des personnes* »¹⁰¹. Dans sa décision Google/DoubleClick, la Commission européenne a indiqué de même que « *la concurrence basée sur la qualité des données collectées n'est pas seulement déterminée par la taille des bases de données respectives, mais également par les différents types de données auxquelles les concurrents ont accès ainsi que par le type qui s'avèrera finalement le plus utile pour la publicité sur*

¹⁰¹ Newman, Nathan, Search, Antitrust and the Economics of the Control of User Data, Yale Journal on Regulation, Vol. 30, No. 3, 2014, également disponible sur SSRN: <http://ssrn.com/abstract=2309547> or <http://dx.doi.org/10.2139/ssrn.2309547>.

¹⁰¹ Executive Office of the President, President's Council of Advisors on Science and Technology, Report to the President - Big Data and Privacy: a Technological Perspective (2014), https://www.whitehouse.gov/sites/default/files/microsites/ostp/PCAST/pcast_big_data_and_privacy_-_may_2014.pdf.

internet » (§ 273). De même, le Information Commissioners' Office, c'est à dire l'autorité indépendante du Royaume-Uni chargée de préserver les droits à l'information dans l'intérêt public, a indiqué que, selon certains praticiens, « *sur les trois V, la variété est la caractéristique la plus important du big data* »¹⁰².

Pour conclure, l'avantage associé à l'accès à un volume de données plus important peut être très différent d'un marché à l'autre. Cela implique une évaluation au cas par cas.

V. Conclusion

L'utilisation des données n'est pas un phénomène nouveau, comme l'illustrent les analyses déjà conduites par les autorités de concurrence à propos de l'avantage retiré de la détention des données sur les marchés traditionnels (non numériques). Conserver des données clients à jour, mener des études de marché ou sonder ses clients constituent depuis longtemps les bases de toute activité commerciale. Cependant, les avancées technologiques et la numérisation de l'économie ont élargi la nature (par ex. les données de géolocalisation alimentées par les smartphones), les sources (par ex. le « tracking » depuis plusieurs appareils pour connaître la navigation d'un internaute), les applications (par ex. « machine learning ») et le volume des données.

L'importance économique des données comme intrant a déjà été prise en compte et, dans plusieurs cas, admise par les autorités de concurrence. A ce jour, le risque de verrouillage du fait de la concentration des données dans les activités numériques a le plus souvent été examiné dans le cadre du contrôle des concentrations. Cela n'exclut pas la mobilisation des outils de répression des pratiques anticoncurrentielles pour appréhender certains comportements en lien avec la collecte et le traitement des données, comme cela a déjà été le cas dans certains marchés traditionnels. Il existe plusieurs types de pratiques liées aux données, relevant de stratégies d'exploitation ou d'éviction, qui pourraient, selon les circonstances de l'espèce, faire l'objet d'une intervention contentieuse.

¹⁰² Information Commissioner's Office, Big data and data protection (2014), <https://ico.org.uk/media/for-organisations/documents/1541/big-data-and-data-protection.pdf> , § 25. Voir également Networked European Software and Services Initiative (NESSI), Big Data: A New World of Opportunities, White Paper (2012), http://www.nessi-europe.com/Files/Private/NESSI_WhitePaper_BigData.pdf.

Cependant, les scénarios anticoncurrentiels qui sous-tendent l'interdiction de telles pratiques supposent, pour la plupart, qu'une entreprise soit en mesure d'asseoir un pouvoir de marché sur la détention d'un ensemble de données que ces concurrents ne peuvent répliquer. Une évaluation spécifique et au cas par cas de la réalité et de l'étendue de l'avantage conféré par les données doit être réalisée pour confirmer ou rejeter ce postulat. Dans cette perspective, il convient de prendre en compte d'abord les caractéristiques propres aux marchés numériques (effets de réseau, multi-hébergement et dynamique concurrentielle) qui se prêtent, ou non, à la création d'un pouvoir de marché, avant de déterminer si les données contribuent à la création ou au renforcement de ce pouvoir de marché.

Deux aspects sont plus particulièrement pertinents lorsqu'il s'agit de déterminer le lien entre données et pouvoir de marché : la rareté et la capacité à reproduire ou à accéder aux données d'une part, le volume et la variété des données d'autre part.

Les données sont « non rivales » au sens où la détention et l'utilisation de données par un opérateur n'excluent pas, en tant que telles, que d'autres opérateurs aient accès aux mêmes données. Le multi-hébergement des consommateurs, de même que la diversification des services offerts par une seule entreprise, offre des opportunités de collectes parallèles de données d'utilisateurs. Cependant, l'accès à ces données peut être conditionné à la capacité d'une entreprise de construire une base de clients suffisamment large, capacité qui dépend à son tour de la mesure dans laquelle les effets de réseaux et d'expérience, tout comme les économies d'échelle, agissent comme des barrières à l'entrée. L'accessibilité à des données de tierces parties, comme les courtiers en données, peut atténuer les problèmes d'accès aux données, mais l'impact d'un recours à ces fournisseurs de données peut varier selon la nature des données concernées, les règles applicables – contractuelles ou réglementaires – protégeant la vie privée de l'utilisateur et la réticence répandue des entreprises à partager leur avantage en matière de données avec leurs concurrents.

La numérisation de l'économie et le déploiement de terminaux connectés captant des activités hors ligne ont considérablement accru la masse de données qu'il est possible de collecter. Ceci a conduit à l'affirmation selon laquelle « les données sont partout » et, en ce sens, ne peuvent pas constituer un intrant susceptible d'être verrouillé. Cependant, cette affirmation ne vaut que si lesdites données sont

réellement accessibles (voir paragraphe précédent). Est également décisive la question de la substituabilité entre les différentes catégories de données – et, en l’absence de substituabilité, celle de la large disponibilité de chaque catégorie de données pour la collecte. Des constats empiriques comme des analyses antérieures de la substituabilité des données par les autorités de concurrence pointent l’existence de différences qui doivent être prises en compte au regard de leur importance dans les affaires en cause (par ex., entre les données mobiles et statiques ; entre les données extraites des requêtes et celles extraites des réseaux sociaux ; entre les données transactionnelles et les données montrant des intentions d’achat qui ne se sont pas concrétisées).

Enfin, l’échelle et la variété des données requises doivent être déterminées. L’importance des données comme intrant stratégique et les opportunités de verrouillage dépendent en partie des volumes de données : (i) auxquels une entreprise peut récolter les bénéfices économiques des données ; (ii) au-delà desquels ces bénéfices déclinent ou cessent complètement d’exister. Ces niveaux varieront selon le type et la fonction des données. Par exemple, la valeur marginale des coordonnées ou des données sociodémographiques peut tendre, dans plusieurs secteurs, à rester à peu près constante (sous réserve des nécessités de mise à jour), tandis que la valeur des données utilisées pour opérer des inférences aura tendance à décroître, au-delà d’un seuil, à mesure que diminue l’impact de chaque observation supplémentaire sur la marge d’erreur. Cependant, un volume significatif et récurrent de données peut devoir être collecté avant d’atteindre ce seuil, par exemple si une activité doit tenir compte, pour rester compétitive, du rythme élevé des changements et des effets de « longue traîne ». La question de la variété des données peut s’avérer aussi importante que celle de l’échelle des données et justifier, selon les conditions de marché et les faits d’espèce, un examen plus approfondi.