

Combien valent nos données ?

La donnée est une matière première pour l'entreprise, à l'origine de la création de valeur. Elle est un levier d'optimisation des ressources. La question de la valeur de cette matière première est insuffisamment traitée dans la littérature économique académique. Elle renvoie au juste équilibre entre protection des données personnelles et utilisation des mêmes données à des fins de valorisation des services proposés par la plateforme, ou à des fins d'innovation.

Une valeur encore mal connue

Abiteboul et Peugeot (2017) distinguent les données qui sont le fruit de l'investissement de l'entreprise, comme le catalogue des produits proposés, et les données qui sont nées des services de l'entreprise et produites par ses clients et ses utilisateurs. Ces données sont coproduites entre une plateforme et ses usagers.

On assimile fréquemment la valeur de la donnée personnelle à la recette publicitaire qu'elle permet. Le tableau ci-dessous fait état du montant des recettes publicitaires des différents médias en France en 2017. Internet domine largement le marché, et les taux de croissance sur l'année 2016-2017 – s'ils se prolongent, montrent que cette domination est appelée à s'accentuer.

Recettes publicitaires des médias, France

	Marché 2017 En M€	Evolution 2016-17
Internet	4094	12%
TV	3286	1%
Presse	2116	-7,4%
Affichage	1180	-2%

Radio	694	-2,6%
Cinéma	95	8,6%

Source IREP.

Ángel et Rubén Cuevas et J. Gonzales (2017) quantifient la valeur publicitaire d'un profil en fonction du temps passé sur Facebook, et montrent que même sans jamais cliquer sur un lien sponsorisé, Facebook génère de la valeur à partir d'un profil. Mais la valorisation des données n'est pas qu'affaire de publicité : Amazon optimise ses prix en fonction des acheteurs grâce à l'exploitation des données qu'elle collecte. Cela permet d'augmenter le volume des ventes.

La tentation est grande pour l'entreprise de stocker un volume important de données, d'autant que les coûts de stockage sont relativement faibles. Elle en espère diverses retombées. En premier lieu, un bon usage des données permet d'améliorer l'efficacité de l'entreprise et de réduire éventuellement les coûts. Il permet par exemple de repérer des possibilités d'améliorer les processus de travail. Le second élément se situe du côté de la demande : meilleure connaissance de la clientèle permettant d'ajuster les services et les biens fournis aux besoins des consommateurs, envoi d'informations ou de publicités ciblées. Il faut ajouter le développement de tout un marché de la donnée : collecte, stockage, analyse. Un troisième élément consiste dans l'individualisation de prestations jusque-là mutualisées : c'est tout le champ de l'assurance, qui pose des questions à la fois éthiques et de politique publique. Ajoutons encore que certaines données individuelles, lorsqu'elles sont traitées en masse et de manière anonymisée, deviennent des biens communs : c'est le cas des données de santé qui ouvrent la voie à des découvertes majeures en matière thérapeutique. Un quatrième terrain de valorisation réside dans l'usage des données par des entreprises étrangères à l'activité de l'entreprise qui les a collectées. Ce pourrait être l'exemple d'un opérateur télécom rassemblant des données de localisation et les proposant à une entreprise de transport ou de tourisme.

Le RGDP et la valeur des données

Nombre de services sont gratuits sur Internet. On est alors dans le cadre de marchés biformes, largement analysés par les économistes depuis les travaux fondateurs de Rochet et Tirole

(2003). La monétisation des données permet de bâtir des modèles économiques dès lors que le service offert est gratuit ou que son coût est très faible.

La tentation est grande pour le consommateur d'effectuer un arbitrage entre le fait d'accepter que ses données soient utilisées à des fins notamment publicitaires et la gratuité. On a souvent évoqué à ce propos l'hypothèse d'une servitude volontaire.

Cela pose plusieurs questions : la première est l'insuffisance de la connaissance de ce que les plates-formes collectent véritablement comme données. On peut faire hypothèse que cette connaissance est d'autant plus faible que les conditions générales d'utilisation sont peu claires et d'une lecture rébarbative. Les coûts de l'information apparaissent alors comme plus élevés, pour l'utilisateur, que les bénéfices à en attendre. La deuxième est la faiblesse de la conscience du prix à payer pour cette gratuité : invasion par la publicité notamment.

Dans tous les cas que nous venons de mentionner plus haut, la question de la protection des données personnelles se pose. Le RGDP (règlement général sur la protection des données personnelles) vient d'être mis en œuvre en France comme dans les autres pays d'Europe. Qu'apporte ce nouveau règlement ? Le texte précise le cadre dans lequel les sociétés pourront exploiter les données personnelles en leur possession (qu'il s'agisse de celles de leurs employés, de leurs clients, ou de leurs fournisseurs). En cas de manquement grave au règlement, l'amende pourra atteindre 20 millions d'euros ou 4 % du chiffre d'affaires mondial de la société. L'importance de la sanction est supposée avoir un effet *ex ante* sur le comportement des entreprises.

En termes économiques, le règlement pose au moins deux séries de questions : la première a trait à ses coûts d'application. S'est d'ailleurs développé un véritable marché destiné à offrir des services de mise en conformité pour les entreprises. La deuxième renvoie à l'équilibre qu'il faut trouver entre protection de la vie privée et développement d'innovations et de services dont le modèle économique repose en large partie sur la valorisation des données.

La tension entre protection des données et exploitation des données est particulièrement forte dans le champ de la santé : si les données de masse permettent de mieux ajuster les mécanismes de prévention, de détection, de soin, lorsqu'elles sont exploitées par des acteurs public ou privés, le risque de revenir à des informations concernant les individus (« désanonymisation ») n'est pas éliminé. La valeur des données de masse est considérable ; selon Abiteboul et Peugeot – op. cité –, on comptait 5000 start-up dans le domaine de l'e-santé en 2016 aux Etats-Unis. Mais la valorisation de ces données pose des questions de

régulation et requiert une vigilance attentive qui ne doit pas être créatrice de barrières à l'innovation.