

Datassets : activation et marketisation des données personnelles

Thomas Beauvisage - SENSE, Orange Labs - thomas.beauvisage@orange.com

Kevin Mellet - SENSE, Orange Labs & CSI, Mines-ParisTech - kevin.mellet@orange.com

Colloque Etienne Thil, Roubaix, octobre 2016

Dans un rapport publié en 2011 intitulé « Personal Data: The Emergence of a New Asset Class », le Forum Economique Mondial développait une vision prospective des données personnelles comme "nouveau 'pétrole' du 21ème siècle", qualificatif largement repris par la suite. Le rapport prévoit notamment que les données personnelles sont appelées à constituer une nouvelle forme de liquidité pour les individus : « En pratique, les données des individus seraient équivalentes à leur "argent". Elles seraient détenues dans un compte où elles seraient contrôlées, gérées, échangées et comptabilisées exactement comme pour les services bancaires aujourd'hui. »

Quelques années plus tard, ce projet semble avoir fait long feu. Certes, les données personnelles ont indéniablement pris une place importante dans les industries numériques, avec la large diffusion des outils de tracking et de personnalisation, notamment à des fins d'action marketing, et l'apparition de véritables courtiers en données personnelles. Pour autant, les internautes semblent avoir peu de prises sur la captation et les usages des informations les concernant, et moins encore de capacité à les monétiser. Cette communication s'attache à dépasser cet apparent paradoxe en décrivant les écosystèmes marchands qui se sont construits autour des données personnelles, et la nature particulière de cet actif pour les annonceurs et les acteurs de l'industrie publicitaire. Elle s'attache pour cela à articuler et croiser deux perspectives sur la place des données dans les marchés. La première s'inscrit dans le sillage d'une tradition critique sur les données issue des Science and Technology Studies, qui s'efforce d'analyser et de contextualiser les configurations marchandes qui se mettent en place autour des infrastructures du big data (Bowker *et al.*, 2009), les nouvelles épistémologies liées à ces outils et ces nouveaux matériaux et traces (Kitchin, 2014), et les outils de personnalisation ordinairement utilisés sur les médias numériques (Turow, 2012; Cardon, 2015). La seconde perspective est celle de la sociologie des marchés. Nous mobilisons ici les travaux ayant montré l'importance des intermédiaires de marché (Bessy et Chauvin, 2013), des investissements de forme (Thevenot, 1986) et des dispositifs de valuation sur les marchés (Muniesa *et al.*, 2007) réalisés par les acteurs pour transformer des ressources en biens et actifs économiques valorisables. Les apports combinés de ces deux champs de recherche nous semblent particulièrement féconds pour aller examiner de manière détaillée la manière dont les données à caractère personnel sont collectées ou fabriquées, transformées en produits et s'insèrent dans des dispositifs plus larges liés aux actions marketing des entreprises.

Sous quelle forme les données personnelles sont-elles mises en marché dans l'univers du marketing et de la publicité ? Pourquoi les internautes sont-ils très largement maintenus à l'écart d'un espace marchand B2B dans lequel les données personnelles s'échangent et se monnaient essentiellement entre professionnels ? Quels sont les acteurs et les enjeux des nouveaux écosystèmes qui se constituent autour de la data ? Bref, comment s'organise la valorisation économique des données à caractère personnel, et quelles sont les perspectives et horizons de ce vaste mouvement de constitution des données comme « actif économique » ?

Nous présentons ici les principaux résultats d'une étude empirique auprès de professionnels, dans les secteurs du marketing et de la publicité¹. Cette communication comprend deux parties. Dans un premier temps, nous étudions les tentatives répétées et infructueuses pour créer une monétisation des données personnelles contrôlée par les internautes. Nous constatons que ce marché n'existe pas, notamment parce que les sommes en jeu sont dérisoires pour le consommateur, et que les industries de la donnée se sont entretemps très largement passées du consentement de l'internaute pour capter ses traces et les valoriser. Dans un deuxième temps, nous nous intéressons au marché B2B des données personnelles. Les activités sur ce marché peuvent être rattachées à deux domaines distincts de la communication des entreprises : le marketing direct (un volet de la communication dite "hors média") et la publicité, dite "communication média". Les spécialistes qui opèrent sur ces marchés n'ont pas attendu la révolution 'big data' pour mettre en place des outils d'agrégation, de traitement et de croisement des données personnelles, faisant de ces dernières une véritable commodité, monnayable et transférable d'une entreprise à l'autre. Ceci dit les transformations en cours ne doivent pas être sous-estimées : c'est

¹ Outre une analyse approfondie de la littérature académique et des publications professionnelles, nous avons participé à une dizaine de conférences professionnelles (Electronic Business Group, Social Media Club, Salon E-Commerce, AdTech Summit) et réalisé des entretiens d'une heure environ avec huit professionnels.

une véritable ingénierie industrielle de la captation, du croisement et de l'exploitation des données individuelles, fortement automatisée, et à très grande échelle, qui se met en place sous nos yeux, avec comme moteur une transformation du produit-même qui est échangé sur le marché – à savoir l'audience – et comme horizon, le rapprochement entre les univers jusqu'ici distincts du marketing direct et de la publicité média.

L'impossible marché C2B des données personnelles

Quelques années après les prévisions du Forum Economique Mondial sur l'émergence des données personnelles comme une nouvelle forme de liquidité pour les citoyens, ce projet semble avoir fait long feu. Les données personnelles ont indéniablement pris une place importante dans les industries numériques : le quotidien des internautes est désormais traversé par la captation d'informations à caractère personnel et leur utilisation à des fins de personnalisation dans la publicité, le e-commerce ou les services. Cependant, alors que les outils de tracking et de calcul du *big data* se sont largement diffusés, et que des courtiers en données personnelles, les *data brokers*, ont émergé, les internautes semblent avoir peu de prises sur la captation et les usages des informations et des traces les concernant, et moins encore de capacité à les monétiser. Qu'est-ce qui explique cet échec d'un marché C2B des données personnelles sous le contrôle des individus ?

L'idée de monétisation des données par les individus connaît en 2010 un intérêt de la part de multiples acteurs. Dans la presse généraliste, une série d'articles soulignent le caractère intrusif des techniques de ciblage publicitaire en ligne et le peu de prises que les internautes ont sur ces systèmes. La thématique de la protection des données personnelles et de leur réappropriation par les internautes émerge comme un axe fort chez les défenseurs des libertés individuelles. Au même moment, chez les fournisseurs de service de l'économie numérique la question de l'interconnexion des services et la mise en place des API laisse entrevoir une plus grande interopérabilité entre services jusqu'ici largement fermés. Dans ce contexte, au croisement des débats et initiatives sur l'*empowerment* de l'internaute et sur le respect de la vie privée (*privacy*) du consommateur numérique, d'un côté, et d'une approche des données personnelles comme patrimoine individuel, de l'autre, se conçoit à cette époque la perspective d'un marché des données personnelles où l'internaute ferait commerce des informations le concernant. Un certain nombre d'acteurs se sont efforcés depuis lors de dresser les contours de ce marché et de lui donner corps.

La valeur des données personnelles au prisme de la vie privée

Comment décrire un marché lorsqu'il n'existe pas encore ? Afin d'esquisser les contours d'un système de vente de leurs données personnelles par les internautes, plusieurs études ont tenté d'estimer la valeur des données personnelles en se plaçant du point de vue de l'offreur (l'internaute).

Ces travaux en économie expérimentale et ces enquêtes ont en commun de projeter l'internaute dans des situations fictives qui l'amènent à mettre en balance la divulgation d'informations le concernant (âge, revenu, historique d'activité sur internet) et un bénéfice financier (réduction, vente des données, etc.). Leur valeur est ainsi envisagée comme le prix – ou son équivalent en termes de service – auquel les individus sont prêts à transférer certaines de leurs données et ce faisant à renoncer à leur *privacy*. Dans cette approche, un premier ensemble de chiffreages s'attache à identifier à la fois les données dont la protection a le plus de valeur pour les individus, les tiers auxquels il serait envisageable pour eux de les céder, et le cas échéant à quel prix (voir tableau). Second point commun de ces études, la valeur est considérée de manière atomique : chaque informations est évaluée seule, et comparée à la valeur attribuée aux autres informations, une somme pouvant dans certains cas constituer une manière d'évaluer la valeur totale des informations d'un individu.

Tableau 1. Synthèse des études sur la valeur des données personnelles pour les individus

Source	Chiffrages
Huberman & al. (2005)	âge = \$57 ; poids = \$74
Beresford & al. (2012)	Couleur préférée et année de naissance = €1
Carrascal & al. (2013)	"information "offline" (age, adresse, statut économique) ≈ €25; historique de navigation = €7 ; interactions sur les réseaux sociaux = €12; historique de <i>search</i> = €2 ; shopping = €5
Staiano & al. (2014)	geolocalisation = €17 à 588 ; communications = €3,4 à 51
Havas Média (2014)	30% vendent dans la tranche maximale, i.e. "plus de 500€"
Orange (2014)	chaque information (nom, numéro de mobile, âge des enfants, revenu, historique d'achats, contacts, etc.) ~ €15 → valeur moyenne de l'ensemble = €170

L'étude menée par Carrascal et al. en 2013 est un bon exemple de ces travaux. L'expérience s'appuie sur un plugin développé pour les navigateurs web, qui demande en temps réel à un panel de

participants (n=168) à quel prix ils seraient prêts à vendre des informations relatives à leur comportement et leurs pratiques sur le web. Les informations 'offline' (âge, adresse, statut économique) ont un prix élevé, autour de 25€ ; le prix moyen de l'historique de navigation est de 7 € avec des variations selon les informations : les interactions sur les réseaux sociaux, 12 € ; la visite de sites financiers, 15,5 € ; le *search*, 2 € ; le *shopping*, 5 €.

Globalement, les prix mentionnés par les internautes dans ce type d'études sont élevés, de l'ordre de quelques euros. Sur ce plan, ces travaux montrent l'importance accordée *a priori* au non-dévoilement de certaines informations les concernant, même si l'on observe une grande disparité de prix, ainsi qu'une dépendance importante des prix aux arbitrages contextuels entre l'accès à un service et le souci de protection des données personnelles. En outre, les conditions pratiques de la cession des données par les individus sont rarement explicitées dans ces tests : s'agit-il d'une vente définitive, de location, où d'une cession à usage unique ? Qui se porterait acquéreur, et pour quelle utilisation ? Surtout, ces mises en situation exploratoires postulent l'existence d'un marché libre des données personnelles au sein duquel les individus auraient le choix de vendre des informations les concernant ; mais ce marché C2B des données personnelles existe-t-il vraiment ?

Des services pleins de promesses

Certaines entreprises ont exploré cette voie, misant sur le respect par les acteurs numériques du consentement des utilisateurs à fournir des informations les concernant, et sur l'attribution de droits de propriété à la donnée personnelle, la constituant ainsi en patrimoine.

Nous avons identifié cinq acteurs dans ce cas, aux Etats-Unis (Personal, Datacoup, Personal Black Box), au Royaume-Uni (Handshake) et en France (Yes Profile). Ces cinq sociétés sont apparues entre 2009 et 2013, au moment même où les industriels du numérique développent des technologies de captation d'informations personnelles fines et systématiques, et où cette captation commence à se constituer en débat dans l'espace public. A la même époque, des artistes ou des performers mettent également leurs traces sur le marché pour questionner la captation et la valorisation marchande de la vie numérique des internautes. Il faut dire que depuis la fin des années 2000, les questions de *tracking* et de *privacy* ponctuent régulièrement le débat public sur les industries numériques. Dans le sillage du rapport aux Etats-Unis de la Federal Trade Commission en 2010 sur l'implémentation du 'Do Not Track' pour limiter le ciblage publicitaire, la question du *tracking* a été largement reprise dans la presse grand public. Quelques années plus tard, le sujet est remis sur la table : en 2013, le Sénat américain publie un rapport détaillé sur les *data brokers*, industriels de la donnée qui collectent et revendent l'accès à d'immenses bases d'informations personnelles. Les conditions semblent donc sur le point d'être réunies pour l'émergence d'un marché donnant aux internautes la liberté de faire le commerce de leurs données personnelles.

Dans ce contexte, nos cinq start-up s'appuient sur un argumentaire commun, que l'on peut résumer ainsi : l'activité économique des grandes sociétés sur internet, est fondée sur la captation de traces dont elles tirent de colossaux bénéfices, notamment par le biais de la publicité. Il s'agit de remettre le consentement de l'internaute au cœur de ces dispositifs et d'en réattribuer les bénéfices aux internautes qui souhaiteraient en faire le commerce. Le modèle de ces sociétés consiste à demander aux internautes de rassembler chez elles les données issues des réseaux sociaux, ainsi que de renseigner un nombre important d'informations les concernant, et notamment celles relatives à leurs intentions d'achat, voire de participer à des focus group et tester des produits. Les gains potentiels pour les internautes participants avancés par ces start-ups sont flatteurs : Handshake avance un bénéfice possible pouvant aller jusqu'à 15.000 £ (19.000 €) annuels par participant ; Yes Profile estime le revenu potentiel mensuel entre 60 € et 100 €, soit jusqu'à 1200 € annuels ; Personal Black Box donne une fourchette comprise entre 50 \$ et 500 \$ (45 et 450 €) ; Datacoup, plus prudent, avance sur son site le chiffre de 8 € par mois, soit une centaine d'euros annuels.

Le marché est ailleurs

Ces services n'ont jamais décollé. En définitive, ces sociétés se heurtent à une difficulté majeure : la captation et la mise en marché des traces et des informations personnelles sur les internautes se fait déjà largement, et sans elles. Etant donné le cadre juridique, les annonceurs n'ont pas besoin de se tourner vers l'internaute pour accéder à des informations qu'elles obtiennent de manière systématique, directement via leurs sites et applications (données des bases CRM, comportemental sur leur site), ou indirectement auprès des acteurs de l'industrie publicitaire et du marketing. Dans ce cas, le consentement de l'utilisateur est soit superflu, soit implicite.

Le fondateur de la société Personal, S. Green, faisait ainsi le constat en 2013 que le marché des données personnelles « n'existe pas pour l'instant parce que les consommateurs ne sont pas de la partie ». On pourrait lui opposer que c'est parce que les consommateurs ne sont pas de la partie que ce marché existe aujourd'hui. Autrement dit, les sociétés qui, comme Yes Profile, promettent aux individus de « redevenir propriétaires de leur profil », ne parviennent pas à tenir cette promesse : le fait de s'inscrire sur le site et d'y déposer ses informations personnelles ne stoppe en rien la captation par d'autres acteurs (régies publicitaires, *ad exchanges*, *data brokers*, etc.), et le profilage opéré par Yes

Profile vient s'ajouter au *tracking* dont les individus sont l'objet, non s'y substituer. En conséquence, le marché C2B de la donnée personnelle n'a jamais vraiment décollé, et les start-ups qui ont tenté de le créer ont une activité des plus réduites – ou ont réorienté leur activité en abandonnant la monétisation de leurs données par les individus (cas de Personal Inc., qui commercialise à présent une solution logicielle collaborative).

Quelques centimes

En outre, quand bien même les internautes reprendraient le contrôle de leurs données et de leurs traces pour en faire le commerce, l'auto-valorisation des données personnelles par les internautes devrait faire face à une autre contrainte de taille : sur les marchés existants, elles ne valent pas grand-chose, et ce prix ne justifie sans doute pas que les particuliers fassent l'effort de les mettre en vente.

Deux sources de calcul s'appuyant sur les tarifs pratiqués sur le marché publicitaire viennent appuyer empiriquement ce constat. Le *Financial Times* a le premier mis en ligne, en juin 2013, un questionnaire interactif permettant aux lecteurs d'évaluer la valeur de certaines de leurs données personnelles (informations sociodémographiques, patrimoine, loisirs et habitudes de consommation). Les évaluations sont basées sur des relevés de prix pratiqués aux Etats-Unis par les courtiers en données personnelles, les *data brokers*. L'outil permet de constater, question après question, le caractère sectoriel de l'intérêt pour les informations individuelles. Dans une perspective marketing, un projet de voyage ou d'achat de voiture, l'arrivée imminente d'un enfant au foyer élèvent la valeur du profil car ils intéressent certains annonceurs en particulier. Et dans le même temps, ces caractéristiques n'incrémentent la valeur totale du profil que de quelques *cents* ou dixièmes de *cents*, constituant une valeur totale moyenne avoisinant les 20 *cents*.

Une seconde étude empirique, issue du laboratoire Privatics de l'INRIA en 2014, apporte un éclairage complémentaire sur la valeur des données personnelles. Dans leur observation des prix d'achat sur les systèmes d'enchères sur le marché publicitaire en ligne (dits « RTB », Real Time Bidding), les auteurs observent que les prix varient en fonction de la profondeur temporelle des informations accumulées sur un internaute par le biais des *cookies*. Ils constatent que les indices d'intentions d'achat suscitent des valorisations plus élevées que le simple profil issu de l'historique de navigation. Cet historique en lui-même n'est valorisé qu'aux environs de 0,0005 \$ seulement pour une impression publicitaire : en rapportant ce prix au nombre moyen de sites visités et de publicités affichées par site, les auteurs évaluent le volume d'affaire généré à 0,18 \$ par mois. La part que les internautes pourraient ainsi prendre de ce marché reviendrait, par an à environ \$0,432, c'est-à-dire quasiment rien.

Ces deux études montrent deux choses : tout d'abord, que la valeur d'un bouquet d'informations personnelles, lorsqu'elle est déterminée par le jeu de l'offre et de la demande sur le marché publicitaire, est dans tous les cas faible, et donc peu attractive en termes de rémunération pour le consommateur. Ensuite, elles lèvent le voile sur le fonctionnement et le type de produits qui définissent le marché de la donnée personnelle aujourd'hui sous l'angle du marketing : un marché B2B des intentions d'achat, des profils qualifiés, des listes d'individus, formaté pour les besoins des annonceurs et d'une chaîne industrielle très automatisée.

Les deux formes de marchandisation des données personnelles pour le marketing

Les tentatives infructueuses de construction d'un marché C2B de la donnée personnelle ont révélé la difficulté, si ce n'est l'impossibilité de permettre aux individus de reprendre le contrôle sur leurs données ; en effet, ces données font l'objet de captations et d'échanges dans des univers marchands dans lesquels le consommateur final n'a pas son mot à dire. Le marché des données est une activité B2B dont la clientèle principale est constituée par les entreprises, et dont les prestataires sont des professionnels établis de longue date, en particulier dans l'univers du marketing et de la publicité. Sous quelle forme les données personnelles sont-elles mises en marché dans l'univers du marketing et de la publicité ? Quels sont les produits constitués et échangés à partir de ce matériau ? Quels sont les acteurs et les enjeux des nouveaux écosystèmes qui se constituent autour de la data ?

Marketing direct : des marchands de liste aux *data brokers*

Les activités de marketing direct des entreprises concernent une communication à destination des consommateurs par le biais de supports considérés comme "directs", par opposition aux supports indirects que sont les médias : traditionnellement, la boîte aux lettres (*mailing*) et le téléphone (*phoning*), et depuis une vingtaine d'années, la boîte mail (*emailing*). Autour de cette activité, s'est constitué un marché dit de la "location d'adresses". Il répond à double enjeu pour les annonceurs : avoir des listes d'adresses, de préférence nominatives et à jour ; et avoir des adresses qualifiées, c'est-à-dire dotées d'informations permettant d'identifier si les individus concernés sont des acheteurs potentiels de leurs produits.

L'histoire de la location d'adresses, qui se développe réellement avec l'informatisation des entreprises dans les années 1970, permet de mieux comprendre la place que les données personnelles y tiennent

aujourd'hui. Voici les éléments majeurs du récit qu'en fait l'un de ses acteurs historiques, ITL, dans son « Petit historique du marché de la location d'adresse », (P. Adam, itl.fr) : « Le début des années soixante vit de nombreuses sociétés se spécialiser dans la fourniture d'adresses par catégories socio professionnelles comme les médecins, dentistes, ingénieurs ou encore élèves de grandes écoles, que l'on regroupait sous le terme éloquent de "personnes à haut pouvoir d'achat". » A la fin des années soixante-dix, l'informatisation des PME apporte un essor à ce marché par le biais de transactions de gré à gré entre entreprises, notamment des acteurs de la VPC (Vente Par Correspondance) « qui se risquent à mettre les noms et adresses de leurs clients à disposition d'autres entreprises contre un volume identique d'adresses ou contre rémunération sur la base d'un prix 'au mille' ».

Comment est-on passé ainsi d'une activité locale et quasi artisanale dans les années soixante aux acteurs internationaux consolidant des bases aux milliers de variables sur des pans larges de populations ? L'informatique constitue la clé de ce changement, qui se joue autant au sein des entreprises clientes de la location d'adresses que chez les marchands de listes et les *data brokers* eux-mêmes. Chez les uns, la numérisation passe non seulement par l'informatisation des bases de CRM et le calcul de scores d'appétence et de fragilité sur ces bases, mais aussi par la généralisation des programmes de fidélisation depuis leur introduction par Dunhumby pour Tesco au Royaume-Uni en 1994. Les entreprises accumulent ainsi une connaissance de leurs clients qui inclut des éléments transactionnels, et sur lesquels elles peuvent travailler dans deux directions : mieux connaître les clients et adapter le discours et l'offre commerciale qui leur est adressée ; et identifier les prospects et les bassins de recrutement de clientèle.

Chez les autres, ce sont essentiellement des mouvements de concentration industrielle qui vont permettre l'émergence des *data brokers* – tels qu'Acxiom, Experian ou Epsilon – qui constituent des mégabases accumulant le plus de variables possibles sur les consommateurs, et les commercialisant sous la forme de fichiers de prospection, de segments marketing décrivant des modes de vie et des habitudes de consommation de manière générale (ex : la segmentation Personix d'Acxiom), de l'enrichissement de bases d'individus existantes, c'est-à-dire l'ajout de variables dans des bases de CRM de leurs clients et de calcul de risque (scoring) sur des individus ou des profils, en particulier dans les secteurs bancaire et assurantiel.

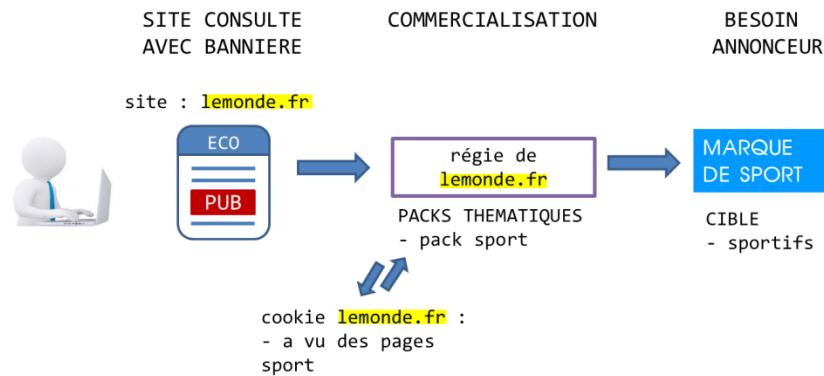
Publicité média : les marchands de cookies

Le marché de la publicité en ligne est l'autre domaine dans lequel les données personnelles font l'objet d'échanges marchands entre acteurs de la chaîne de valeur publicitaire. Dans cet univers, la clé de voûte des échanges économiques est l'audience, qui joue le rôle de monnaie d'échange entre les différents acteurs du secteur : les annonceurs achètent aux médias l'accès à des audiences, dont les instituts de mesure (tels que Médiamétrie ou Nielsen) permettent de mesurer les caractéristiques, et donc la valeur.

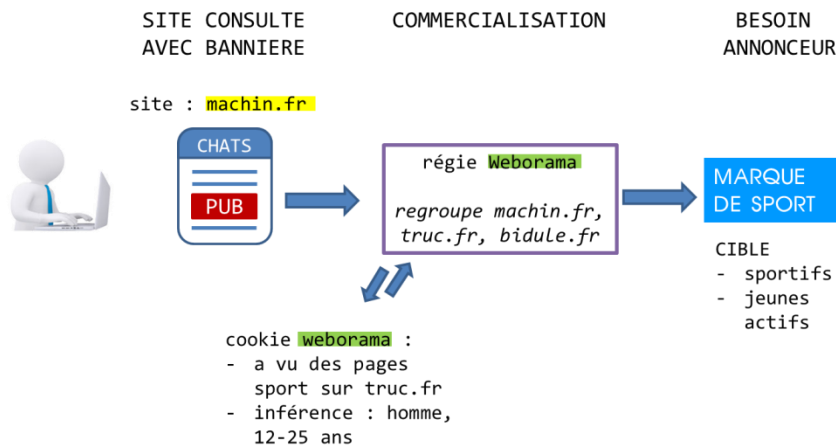
Toute l'évolution du secteur de la publicité en ligne depuis sa création s'articule autour de la prise en compte de plus en plus large des comportements de navigation des internautes pour qualifier et estimer la valeur des espaces publicitaires, créer des produits au plus près de ce qui s'apparente à des intentions d'achats, et faire commerce de ces informations de profil entre acteurs de la chaîne de valeur. C'est à une redéfinition du produit-audience que l'on a progressivement assisté, et dont le pivot est le cookie de navigation.

Le cookie est un mécanisme qui permet à un site web de stocker des informations sur les navigateurs qui communiquent avec lui. Les cookies sont très utiles pour la navigation, puisqu'ils servent à ne pas "redécouvrir" un internaute inconnu à chaque page qu'il visite sur un site : ils permettent d'avoir des "sessions" de navigation, de remplir des paniers, de sauvegarder des préférences, etc.

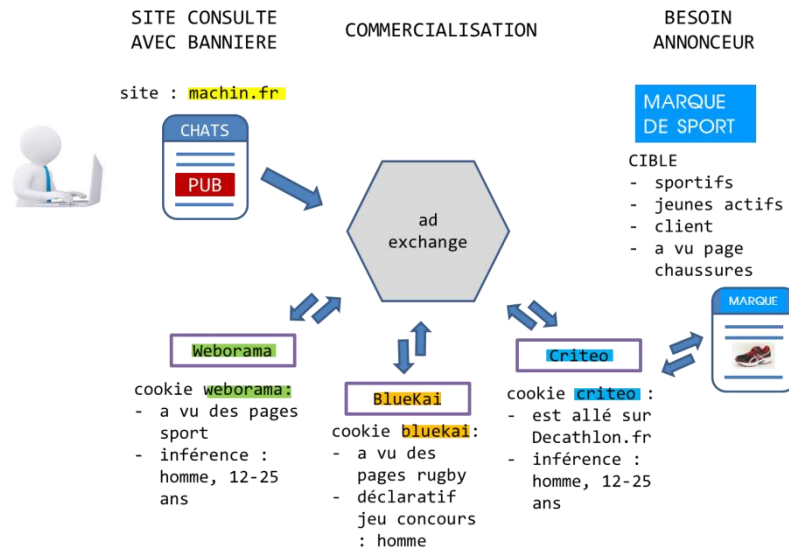
Dans les premiers temps de la publicité en ligne, la tarification et la vente se font sur le modèle des médias traditionnels : c'est la mesure d'audience qui sert à qualifier le profil des visiteurs d'un site, à quoi les régies ajoutent des qualifications par le contenu des pages consultées. La première évolution que le cookie apporte à ce modèle est la possibilité pour le site et sa régie de découpler la visite de pages sur le thème "sport" de la vente de bannières liées au sport : en traçant la navigation de ses visiteurs dans la durée, le site peut satisfaire la demande d'un annonceur de cibler des amateurs de sport à n'importe quel endroit du site.



Ce premier découplage entre la qualification du visiteur et les contenus consultés est vite doublé d'un second à plus large échelle : des régies multi-sites (ou *ad networks*) comme Weborama agrègent la connaissance des parcours des internautes sur l'ensemble des sites partenaires – qui autorisent Weborama à déposer puis enrichir des cookies sur leurs visiteurs. Dès lors, un annonceur cherchant des amateurs de sport peut les trouver sur un site affilié à Weborama dont les contenus n'ont aucun rapport avec le sport, il suffit que l'internaute ait visité un autre site partenaire qui soit lié à la thématique "sport".



La seconde évolution majeure de ce marché est le développement de la *publicité programmatique*. Le programmatique repose sur la commercialisation des espaces publicitaires par l'intermédiaire de places de marché (ou *ad exchanges*), qui s'appuient sur des mécanismes d'enchères réalisées le plus souvent en quasi temps réel, d'où le qualificatif de *real-time bidding* (RTB). Ce mécanisme de coordination impose un traitement totalement automatisé : les stratégies des acheteurs et vendeurs sont mises en œuvre par des algorithmes qui évaluent l'opportunité d'afficher une bannière à chaque impression et cherchent à en optimiser la valeur. Au sein des *ad exchanges*, les espaces publicitaires mis aux enchères sont sujets à de fortes asymétries d'information : certains acteurs disposent d'informations particulières, d'autres non, et c'est le jeu de l'offre et de la demande qui contribue à fixer le prix final des espaces publicitaires.



L'élément central est que ces places de marché sont intrinsèquement conçues pour la mise en concurrence des acheteurs sur la base des informations de cookie dont ils disposent. En effet, le système d'enchères permet à tous les enchérisseurs de participer aux enchères en mobilisant l'ensemble des informations qu'ils détiennent sur l'internaute, même s'ils n'ont aucun lien avec le site qu'il est en train de visiter. Par conséquent, les différents participants aux enchères – acheteurs finaux (annonceurs et agences) ; courtiers qui achètent et revendent aussitôt des espaces publicitaires après les avoir requalifiés (ad networks, data providers) ; vendeurs (éditeurs) – doivent construire leur connaissance en amont de leur participation aux enchères.

Les différents protagonistes de la chaîne de valeur programmatique, qu'ils soient acheteurs finaux, vendeurs, ou intermédiaires, peuvent s'appuyer, pour construire leur connaissance, soit sur les informations dont ils disposent, soit sur des informations acquises auprès de fournisseurs de données : les *data providers*. Les *data providers* collectent, enrichissent, combinent et revendent des cookies. L'enjeu principal pour ces acteurs est de couvrir une large surface d'usage sur le web pour observer les comportements des internautes sur une grande variété de sites et une grande part de leurs usages du web. Pour ce faire, les *data providers* nouent des partenariats avec différents acteurs disposant de gisements de cookies (éditeurs et autres fournisseurs de données), auprès desquels ils se fournissent en données individuelles moyennant des accords de mutualisation de données de cookies, de partage de connaissance, ou de partage des revenus. Les *data providers* sont donc autant des collecteurs d'informations individuelles que des intégrateurs de bases dont la plupart sont souvent sectorielles (voiture, banque, intentions d'achat, etc.). Les principaux acteurs sur ce segment sont des pure players web – Exelate, BlueKai, Weborama, Ezakus, Nugg-Ad, etc – mais il est à noter que certains *data brokers* qui venaient des bases de données offline se sont également positionnés sur cette activité (par exemple : Acxiom ou Experian).

Les *data providers* sont finalement parvenus à construire une offre qui s'insère techniquement et commercialement dans la chaîne de valeur de la publicité programmatique.

Vers une convergence du marketing direct et de la publicité média ?

La donnée personnelle intervient donc dans le marketing de deux manières, historiquement distinctes : sous forme de coordonnées et de caractéristiques sociodémographiques stables et (plutôt) vérifiées dans l'univers du marketing direct, et sous forme de cookies agrégeant des traces de navigation et des caractéristiques sociodémographiques (plutôt) prédites dans l'univers de la publicité. Ces deux univers sont appelés aujourd'hui à converger, notamment avec la mise en place de *Data Management Platforms* (DMP) chez les annonceurs.

En termes simples, une DMP est une base de données permettant de rassembler, de trier et d'exploiter des données provenant de sources hétérogènes. L'offre de logiciels DMP a émergé à partir de 2011 et s'est développée rapidement depuis, sur la base d'une triple promesse :

- i) rassembler dans une même "bassine à data" des données de contacts provenant de sources hétérogènes ;
- ii) opérer des traitements statistiques (segmentation, scoring, etc) sur ces données agrégées ;
- iii) mettre en place des actions marketing variées sur les segments d'audience "sur mesure" ainsi constitués.

Le déploiement des DMP s'inscrit dans la perspective du "marketing relationnel", qui consiste pour l'entreprise à développer dans la durée une relation individualisée avec ses différentes audiences en

fonction de leur proximité-éloignement : client et prospect (les ancrages du CRM), suspect et anonyme (les ancrages de la publicité média).

L'objectif est ambitieux, et soulève un certain nombre de difficultés d'ordre technique (et réglementaire) : embarquement ("onboarding") des données clients dans la DMP, fusion de ces données avec des données de navigation internes ou issues de tiers, exploitation et activation en temps réel de ces données dans la chaîne de valeur programmatique. Ces opérations d'alimentation et d'exploitation de la base de données, dont la difficulté est souvent minimisée par les fournisseurs de solutions de DMP, sont encore largement le fruit d'un travail manuel et d'une certaine forme de "bricolage" de la data. La mise en place des DMP relève encore de l'expérimentation et la promesse globale de réconciliation des univers du marketing direct et de la publicité média, autour de la donnée personnelle, devra faire l'objet d'investigations et d'évaluations.

Perspectives : quel statut pour les données personnelles sur les marchés ?

En essayant de suivre concrètement et techniquement la manière dont les données à caractère personnel sont collectées et constituées en tant que produits, nous sommes arrivés dans un premier temps à constater que ce type d'information ne constitue pas une forme de liquidité pour les individus, contrairement à ce que le Forum Economique Mondial anticipait. Les données personnelles font l'objet de transactions par des acteurs spécialisés. En ce sens, les données pourraient apparaître comme une forme de commodité au sein du secteur de la publicité en ligne. Pour autant, ce premier constat mérite d'être dépassé.

En premier lieu, les données personnelles, qu'elles soient sous forme de listes de contact ou sous forme de cookies encapsulant des connaissances comportementales, prennent place dans des architectures automatisées de la publicité en ligne où elles servent avant tout à valoriser les espaces publicitaires. En ce sens, les données s'apparentent à une forme renouvelée de *currency* (Napoli, 2003) pour estimer et monétiser les contacts des annonceurs avec leurs acheteurs potentiels, que ce soit des espaces publicitaires ou des listes d'adresses pour le marketing direct. Dans cette perspective, les données sont des outils de coordination et de valuation (Muniesa *et al.*, 2007).

Par ailleurs, la démocratisation des outils de traçage des consommateurs sur les supports numériques, ainsi que la constitution de bases de CRM orientées vers une historicisation des transactions (à quoi s'ajoute la possibilité de compléter ces bases avec les informations fournies par les data brokers), renouvelle plus profondément le statut des données à caractère personnels. Pour les annonceurs, à la différence des médias et des acteurs intermédiaires de l'industrie publicitaire, l'information accumulée devient aujourd'hui un levier spécifique permettant plus d'efficacité dans les actions de marketing (ciblage, scoring, onboarding) qui fait l'objet d'investissements dans des architectures informatiques, dans des processus et dans des structures organisationnelles spécifiques. En ce sens, la donnée devient, entre les mains des annonceurs une nouvelle catégorie d'actifs, des *datassets*. Mais cet actif, loin de faire l'objet d'une circulation et d'une marchandisation fluide, s'insère dans des stratégies de captation et de territorialisation des zones d'observations qui soulignent la nature ambiguë de ce bien informationnel.

Bibliographie

Beresford, A. R., Kübler, D., and Preibusch, S. (2012), "Unwillingness to pay for privacy: A field experiment". *Economics Letters*, 117(1), 25-27.

Bessy C., Chauvin P. M. (2013). "The power of market intermediaries: From information to valuation processes". *Valuation studies*, 1(1), 83-117.

Bowker G. C., Baker K., Millerand F., Ribes D. (2009). "Toward information infrastructure studies: Ways of knowing in a networked environment". In *International handbook of internet research* (pp. 97-117). Springer Netherlands.

Cardon D. (2015). *A quoi rêvent les algorithmes: Nos vies à l'heure des big data*. Seuil.

Carrascal J. P., Riederer C., Erramilli V., Cherubini M., de Oliveira R. (2013). "Your browsing behavior for a big mac: Economics of personal information online", *Proceedings of WWW'2013*, ACM.

Huberman B.A., Adar E., Fine L.R. (2005). "Valuating privacy". *IEEE Security & Privacy*, 3(5), 22-25.

Kitchin R. (2014). "Big Data, new epistemologies and paradigm shifts". *Big Data & Society*, 1 (1).

Muniesa F., Millo Y., Callon M. (2007). "An introduction to market devices". *The sociological review*, 55 (s2), 1-12.

Napoli P. M. (2003) *Audience economics: Media institutions and the audience marketplace*. Columbia University Press.

Staiano J., Oliver N., Lepri B., de Oliveira R., Caraviello M., Sebe N. (2014), "Money walks: a human-centric study on the economics of personal mobile data", *Proceedings of the 2014 ACM International Joint Conference on Pervasive and Ubiquitous Computing*, ACM

Thévenot L. (1986). "Les investissements de forme". *Conventions économiques*, 29, 21-71.

Turow J. (2012). *The daily you: How the new advertising industry is defining your identity and your worth*. Yale University Press.

US Senate (2013). *A Review of the Data Broker Industry: Collection, Use and Sale of Consumer Data for Marketing Purpose*, Staff report for chairman Rockefeller.