

Comment les entreprises surveillent notre quotidien

Vous croyez tout savoir déjà sur l'exploitation de nos données personnelles ? Parcourez plutôt quelques paragraphes de ce très vaste dossier...

Il s'agit du remarquable travail d'enquête procuré par Cracked Labs, une organisation sans but lucratif qui se caractérise ainsi :

... un institut de recherche indépendant et un laboratoire de création basé à Vienne, en Autriche. Il étudie les impacts socioculturels des technologies de l'information et développe des innovations sociales dans le domaine de la culture numérique.

... Il a été créé en 2012 pour développer l'utilisation participative des technologies de l'information et de la communication, ainsi que le libre accès au savoir et à l'information - indépendamment des intérêts commerciaux ou gouvernementaux. Cracked Labs se compose d'un réseau interdisciplinaire et international d'experts dans les domaines de la science, de la théorie, de l'activisme, de la technologie, de l'art, du design et de l'éducation et coopère avec des parties publiques et privées.

Bien sûr, vous connaissez les GAFAM omniprésents aux avant-postes pour nous engluer au point que s'en déprendre complètement est difficile... Mais connaissez-vous Acxiom et LiveRamp, Equifax, Oracle, Experian et TransUnion ? Non ? Pourtant il y a des chances qu'ils nous connaissent bien...

Il existe une industrie très rentable et très performante des données « client ».

Dans ce long article documenté et qui déploie une vaste gamme d'exemples dans tous les domaines, vous ferez connaissance avec les coulisses de cette industrie intrusive pour laquelle il semble presque impossible de « passer inaperçu », où notre personnalité devient un profil anonyme mais tellement riche de renseignements que nos nom et prénom n'ont aucun intérêt particulier.

L'article est long, vous pouvez préférer le lire à votre rythme en format .PDF (2,3 Mo)

-> framablog.org-Comment-les-entreprises-surveillent-notre-quotidien-NEW

L'équipe de Framalang s'est largement mobilisée pour vous procurer cette longue traduction : Abel, mo, Moutmout, Penguin, Opsylac, Luc, Lyn., hello, Jérochat, QS, Jérochat, Asta, Mannik, roptat, audionuma, Opsylac, Lumibd, linuxmario, goofy et un anonyme.

Des entreprises mettent notre quotidien sous surveillance

Source : <http://crackedlabs.org/en/corporate-surveillance>

Par Wolfie Christl

avec les contributions de : Katharina Kopp, Patrick Urs Riechert / Illustrations de Pascale Osterwalder.

Comment des milliers d'entreprises surveillent, analysent et influencent la vie de milliards de personnes. Quels sont les principaux acteurs du pistage numérique aujourd'hui ? Que peuvent-ils déduire de nos achats, de nos appels téléphoniques, de nos recherches sur le Web, de nos *Like* sur Facebook ? Comment les plateformes en ligne, les entreprises technologiques et les courtiers en données font-ils pour collecter, commercialiser et exploiter nos données personnelles ?

Ces dernières années, des entreprises dans de nombreux secteurs se sont mises à surveiller, pister et suivre les gens dans **pratiquement tous les aspects de leur vie**. les comportements, les déplacements, les relations sociales, les centres d'intérêt, les faiblesses et les moments les plus intimes de milliards de personnes sont désormais continuellement enregistrés, évalués et analysés en temps réel. L'exploitation des données personnelles est devenue une industrie pesant plusieurs milliards de dollars. Pourtant, de ce pistage numérique omniprésent, on ne voit que **la partie émergée de l'iceberg** ; la majeure partie du processus se déroule dans les coulisses et reste opaque pour la plupart d'entre nous.

Ce rapport de Cracked Labs examine le fonctionnement interne et les pratiques en vigueur dans cette industrie des données personnelles. S'appuyant sur des années de recherche et sur un précédent rapport de 2016, l'enquête donne à voir la circulation cachée des données entre les entreprises. Elle cartographie la

structure et l'étendue de **l'écosystème numérique de pistage et de profilage** et explore tout ce qui s'y rapporte : les technologies, les plateformes, les matériels ainsi que les dernières évolutions marquantes.

Le rapport complet (93 pages, en anglais) est disponible en téléchargement au format PDF, et cette publication web en présente un résumé **en dix parties**.

Sommaire

- I. Analyser les individus
- II. Analyser les individus dans la finance, les assurances et la santé
- III. Collecte et utilisation massives de données client
- IV. Les courtiers en données/ et le marché des données personnelles
- V. La surveillance en temps réel des comportements quotidiens
- VI. Relier, faire correspondre et combiner des profils numériques
- VII. Gérer les clients et les comportements : personnalisation et contrôle
- VIII. Dans les mailles du filet : vie quotidienne, données commerciales et analyse du risque
- IX. Cartographie de l'écosystème du pistage et du profilage commercial
- X. Vers une société du contrôle social numérique généralisé ?

En 2007, Apple a lancé le smartphone, Facebook a atteint les 30 millions d'utilisateurs, et des entreprises de publicité en ligne ont commencé à cibler les internautes en se basant sur des données relatives à leurs préférences individuelles et leurs centres d'intérêt. Dix ans plus tard, **un large ensemble d'entreprises dont le cœur de métier est les données (les data-companies ou entreprises de données en français) a émergé**, on y trouve de très gros acteurs comme Facebook ou Google mais aussi des milliers d'autres entreprises, qui sans cesse, **se partagent et se vendent les unes aux autres des profils numériques**. Certaines entreprises ont commencé à combiner et à relier des données du web et des smartphones avec les données clients et les informations hors-ligne qu'elles avaient accumulées pendant des décennies.

La machine omniprésente de surveillance en temps réel qui a été développée pour la publicité en ligne s'étend rapidement à d'autres domaines, **de la tarification à la communication politique en passant par le calcul de solvabilité et la gestion des risques**. Des plateformes en ligne énormes, des entreprises de publicité numérique, des courtiers en données et des entreprises de divers

secteurs peuvent maintenant identifier, trier, catégoriser, analyser, évaluer et classer les utilisateurs via les plateformes et les matériels. **Chaque clic sur un site web** et chaque mouvement du doigt sur un smartphone peut activer un large éventail de mécanismes de partage de données distribuées entre plusieurs entreprises, ce qui, en définitive, **affecte directement les choix offerts aux gens**. Le pistage numérique et le profilage, en plus de la personnalisation ne sont pas seulement utilisés pour surveiller, mais aussi pour influencer les comportements des personnes.

Vous devez vous battre pour votre vie privée, sinon vous la perdrez.

Eric Schmidt, Google/Alphabet, 2013

Analyser les individus

Des études scientifiques démontrent que de nombreux aspects de la personnalité des individus peuvent être déduits des données générées par des recherches sur Internet, des historiques de navigation, des comportements lors du visionnage d'une vidéo, des activités sur les médias sociaux ou des achats. Par exemple, des **données personnelles sensibles** telles que l'origine ethnique, les convictions religieuses ou politiques, la situation amoureuse, l'orientation sexuelle, ou l'usage d'alcool, de cigarettes ou de drogues peuvent être assez précisément déduites des **Like sur Facebook d'une personne**. L'analyse des profils de réseaux sociaux peut aussi prédire des traits de personnalité comme la stabilité émotionnelle, la satisfaction individuelle, l'impulsivité, la dépression et l'intérêt pour le sensationnel.

Analyser les *like* Facebook, les données des téléphones, et les styles de frappe au clavier



Pour plus de détails, se référer à Christl and Spiekermann 2016 (p.

14-20). Sources : Kosinski et al 2013, Chittaranjan et al 2011, Epp et al 2011.

De la même façon, il est possible de déduire certains traits de caractères d'une personne à partir de données sur les sites Web qu'elle a visités, sur les appels téléphoniques qu'elle a passés, et sur les applis qu'elle a utilisées. **L'historique de navigation** peut donner des informations sur la profession et le niveau d'étude. Des chercheurs canadiens ont même réussi à évaluer des états émotionnels comme la confiance, la nervosité, la tristesse ou la fatigue en analysant la **façon dont on tape sur le clavier de l'ordinateur**.

Analyser les individus dans la finance, les assurances et la santé

Les résultats des méthodes actuelles d'extraction et d'analyse des données reposent sur des **corrélations statistiques** avec un certain niveau de probabilité. Bien qu'ils soient significativement plus fiables que le hasard dans la prédiction des caractéristiques ou des traits de caractère d'un individu, ils ne sont évidemment pas toujours exacts. Néanmoins, ces méthodes sont déjà mises en œuvre pour **trier, catégoriser, étiqueter, évaluer, noter et classer les personnes**, non seulement dans une approche marketing mais aussi pour prendre des décisions dans des domaines riches en conséquence comme la finance, l'assurance, la santé, pour ne citer qu'eux.

L'évaluation de crédit basée sur les données de comportement numérique

Des startups comme Lenddo, Kreditech, Cignifi et ZestFinance utilisent déjà les données récoltées sur les réseaux sociaux, lors de recherches sur le web ou sur les téléphones portables pour **calculer la solvabilité d'une personne** sans même utiliser de données financières. D'autres se basent sur la façon dont quelqu'un va remplir un formulaire en ligne ou naviguer sur un site web, sur la grammaire et la ponctuation de ses textos, ou sur l'état de la batterie de son téléphone. Certaines entreprises incluent même des données sur les amis avec lesquels une personne est connectée sur un réseau social pour évaluer sa

solvabilité.

Cignifi, qui calcule la solvabilité des clients en fonction **des horaires et de la fréquence des appels téléphoniques**, se présente comme « la plateforme ultime de monétisation des données pour les opérateurs de réseaux mobiles ». De grandes entreprises, notamment MasterCard, le fournisseur d'accès mobile Telefonica, les agences d'évaluation de solvabilité Experian et Equifax, ainsi que le géant chinois de la recherche web Baidu, ont commencé à nouer des partenariats avec des startups de ce genre. L'application à plus grande échelle de services de cette nature est particulièrement en croissance dans les pays du Sud, ainsi qu'auprès de groupes de population vulnérables dans d'autres régions.

Réciproquement, les données de crédit nourrissent le marketing en ligne. Sur Twitter, par exemple, les annonceurs peuvent cibler leurs publicités en fonction de la **solvabilité supposée des utilisateurs de Twitter** sur la base des données client fournies par le courtier en données Oracle. Allant encore plus loin dans cette logique, **Facebook** a déposé un brevet pour une évaluation de crédit basée sur la cote de solvabilité de vos amis sur un réseau social. Personne ne sait s'ils ont l'intention de réellement mettre en application cette intégration totale des réseaux sociaux, du marketing et de l'évaluation des risques.

On peut dire que toutes les données sont des données sur le crédit, mais il manque encore la façon de les utiliser.

Douglas Merrill, fondateur de ZestFinance et ancien directeur des systèmes d'informations chez Google, 2012

Prédire l'état de santé à partir des données client

Les entreprises de données et les assureurs travaillent sur des programmes qui utilisent les informations sur la vie quotidienne des consommateurs pour prédire leurs risques de santé. Par exemple, l'assureur **Aviva**, en coopération avec la société de conseil Deloitte, a utilisé des données clients achetées à un courtier en données et habituellement utilisées pour le marketing, pour prédire les **risques de santé individuels** (comme le diabète, le cancer, l'hypertension et la dépression) de 60 000 personnes souhaitant souscrire une assurance.

La société de conseil McKinsey a aidé à **prédire les coûts hospitaliers** de

patients en se basant sur les données clients d'une « grande compagnie d'assurance » santé américaine. En utilisant les informations concernant la démographie, la structure familiale, les achats, la possession d'une voiture et d'autres données, McKinsey a déclaré que ces « renseignements peuvent aider à identifier des sous-groupes stratégiques de patients avant que des périodes de coûts élevés ne surviennent ».

L'entreprise d'analyse santé **GNS Healthcare** a aussi calculé les risques individuels de santé de patients à partir d'un large champ de données tel que la génétique, les dossiers médicaux, les analyses de laboratoire, les appareils de santé mobiles et le comportement du consommateur. Les sociétés partenaires des assureurs tels que Aetna donnent une note qui identifie « les personnes susceptibles de subir une opération » et proposent de **prédire l'évolution de la maladie et les résultats des interventions**. D'après un rapport sectoriel, l'entreprise « classe les patients suivant le retour sur investissement » que l'assureur peut espérer s'il les cible pour des interventions particulières.

LexisNexis Risk Solutions, à la fois, un important courtier en données et une société d'analyse de risque, fournit un produit d'évaluation de santé qui calcule les risques médicaux ainsi que les frais de santé attendus individuellement, en se basant sur une importante quantité de données consommateurs, incluant les achats.

Collecte et utilisation massives de données client

Les plus importantes plates-formes connectées d'aujourd'hui, **Google et Facebook** en premier lieu, ont des informations détaillées sur la vie quotidienne de milliards de personnes dans le monde. Ils sont les plus visibles, les plus envahissants et, hormis les entreprises de renseignement, les publicitaires en ligne et les services de détection des fraudes numériques, peut-être les acteurs les plus avancés de **l'industrie de l'analyse et des données personnelles**. Beaucoup d'autres agissent en coulisse et hors de vue du public.

Le cœur de métier de la **publicité en ligne** consiste en un écosystème de milliers d'entreprises concentrées sur la traque constante et le profilage de milliards de personnes. À chaque fois qu'une publicité est affichée sur un site web ou une application mobile, un profil d'utilisateur vient juste d'être vendu au plus gros enchérisseur dans les millisecondes précédentes. Contrairement à ces nouvelles

pratiques, les **agences d'analyse de solvabilité** et les **courtiers en données clients** exploitent des données personnelles depuis des décennies. Ces dernières années, ils ont commencé à combiner les très nombreuses données dont ils disposent sur la vie hors-ligne des personnes avec les bases de données utilisateurs et clients utilisées par de grandes plateformes, par des entreprises de publicité et par une multitude d'autres entreprises dans de nombreuses secteurs.

Les entreprises de données ont des informations détaillées sur des milliards de personnes

Plateformes en ligne grand public

Facebook dispose des profils de	1,9 milliards d'utilisateurs de Facebook 1,2 milliards d'utilisateurs de Whatsapp 600 millions d'utilisateurs d'Instagram
Google dispose des profils de	2 milliards d'utilisateurs d'Android + d'un milliard d'utilisateurs de Gmail + d'un milliard d'utilisateurs de Youtube
Apple dispose des profils de	1 milliard d'utilisateurs d'iOS

Sociétés d'analyse de la solvabilité

Experian	dispose des données de solvabilité de 918 millions de personnes dispose des données marketing de 700 millions de personnes a un "aperçu" sur 2,3 milliards de personnes
Equifax	dispose des données de 820 millions de personnes et d'1 milliard d'appareils
TransUnion	dispose des données d'1 milliard de personnes

Courtiers en données clients

Acxiom	dispose des données de	700 millions de personnes
		1 milliard de cookies et d'appareils mobiles
		3,7 milliards de profils clients
Oracle	dispose des données de	1 milliard d'utilisateurs d'appareils mobiles
		1,7 milliards d'internautes
	donne accès à	5 milliards d'identifiants uniques client

Facebook utilise au moins 52 000 caractéristiques personnelles pour trier et classer ses 1,9 milliard d'utilisateurs suivant, par exemple, leur orientation politique, leur origine ethnique et leurs revenus. Pour ce faire, la plateforme analyse leurs messages, leurs *Likes*, leurs partages, leurs amis, leurs photos, leurs mouvements et beaucoup d'autres comportements. De plus, Facebook **acquiert à d'autres entreprises des données sur ses utilisateurs**. En 2013, la plateforme démarre son partenariat avec les quatre courtiers en données Acxiom, Epsilon, Datalogix et BlueKai, les deux derniers ont ensuite été rachetés par le géant de l'informatique Oracle. Ces sociétés aident Facebook à **pister et profiler** ses utilisateurs bien mieux qu'il le faisait déjà en lui fournissant des données collectées en dehors de sa plateforme.

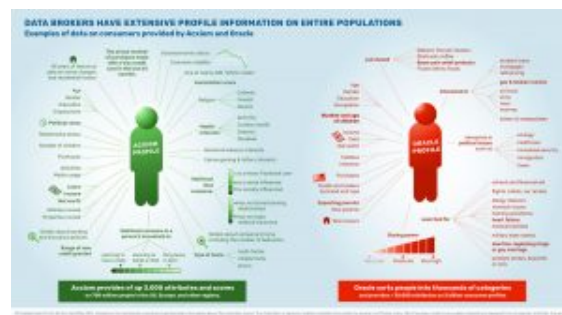
Les courtiers en données et le marché des données personnelles

Les courtiers en données client ont un rôle clé dans le marché des données personnelles actuel. Ils agrègent, combinent et échangent des quantités astronomiques d'informations sur des populations entières, collectées depuis des sources en ligne et hors-ligne. Les courtiers en données collectent de

l'information disponible publiquement et achètent le droit d'utiliser les données clients d'autres entreprises. Leurs données proviennent en général de **sources qui ne sont pas les individus eux-mêmes**, et sont collectées en grande partie sans que le consommateur soit au courant. Ils analysent les données, en font des déductions, construisent des catégories de personnes et fournissent à leurs clients des informations sur des milliers de caractéristiques par individu.

Dans les profils individuels créés par les courtiers en données, on trouve non seulement des informations à propos de l'éducation, de l'emploi, des enfants, de la religion, de l'origine ethnique, de la position politique, des loisirs, des centres d'intérêts et de l'usage des médias, mais aussi à propos du comportement en ligne, par exemple les recherches sur Internet. Sont également collectées les données sur les achats, l'usage de carte bancaire, le revenu et l'endettement, la gestion bancaire et les polices d'assurance, la propriété immobilière et automobile, et tout un tas d'autres types d'information. Les courtiers en données calculent et **attribuent aussi des notes aux individus afin de prédire leur comportement futur**, par exemple en termes de stabilité économique, de projet de grossesse ou de changement d'emploi.

Quelques exemples de données clients fournies par Acxiom et Oracle



exemples de données clients
fournies par Acxiom et Oracle (en
avril/mai 2017) - sources : voir le
rapport

Acxiom, un important courtier en données

Fondée en 1969, Acxiom gère l'une des plus grandes bases de données client commerciales au monde. Disposant de milliers de sources, l'entreprise fournit jusqu'à 3000 types de données sur 700 millions de personnes réparties dans de

nombreux pays, dont les États-Unis, le Royaume-Uni et l'Allemagne. Née sous la forme d'une entreprise de marketing direct, Acxiom a développé ses bases de données client centralisées à la fin des années 1990.

À l'aide de son système **Abilitek Link**, l'entreprise tient à jour une sorte de registre de la population dans lequel chaque personne, chaque foyer et chaque bâtiment reçoit un identifiant unique. En permanence, l'entreprise met à jour ses bases de données sur la base d'informations concernant **les naissances et les décès, les mariages et les divorces, les changements de nom ou d'adresse** et aussi bien sûr de nombreuses autres données de profil. Quand on lui demande des renseignements sur une personne, Acxiom peut par exemple donner une appartenance religieuse parmi l'une des 13 retenues comme « catholique », « juif », ou « musulman » et une appartenance ethnique sur quasiment 200 possibles.

Acxiom commercialise l'accès aux profils détaillés des consommateurs et aide ses clients à trouver, cibler, identifier, analyser, trier, noter et classer les gens. L'entreprise gère aussi directement pour ses propres clients **15 000 bases de données clients** représentant des milliards de profils consommateurs. Les clients d'Acxiom sont des grandes banques, des assureurs, des services de santé et des organismes gouvernementaux. En plus de son activité de commercialisation de données, Acxiom fournit également des services de vérification d'identité, de gestion du risque et de détection de fraude.

Acxiom et ses fournisseurs de données, ses partenaires et ses services



Acxiom et ses fournisseurs de données, ses partenaires et ses clients (en avril/mai 2017) - sources : voir le rapport

Depuis l'acquisition en 2014 de la société de données en ligne **LiveRamp**, Acxiom a déployé d'importants efforts pour connecter son dépôt de données - couvrant une dizaine d'années - au monde numérique. Par exemple, Acxiom était parmi les premiers courtiers en données à fournir de l'information additionnelle à Facebook, Google et Twitter afin d'aider ces plateformes à mieux pister ou catégoriser les utilisateurs en fonction de leurs achats mais aussi en fonction d'autres comportements qu'ils ne savaient pas encore eux-mêmes pister.

LiveRamp de Acxiom **connecte et combine les profils numériques** issus de centaines d'entreprises de données et de publicité. Au centre se trouve son système IdentityLink, qui aide à reconnaître les individus et à relier les informations les concernant, dans les bases de données, les plateformes et les appareils en se basant sur leur adresse de courriel, leur numéro de téléphone, l'identifiant de leur téléphone, ou d'autres identifiants. Bien que l'entreprise assure que les correspondances et les associations se fassent de manière « anonyme » et « dé-identifiée », elle dit aussi pouvoir « connecter des données hors-ligne et en ligne sur un seul identifiant ».

Parmi les entreprises qui ont récemment été reconnues comme étant des fournisseurs de données par LiveRamp, on trouve les géants de l'analyse de solvabilité **Equifax, Experian et TransUnion**. De plus, de nombreux services de pistage numérique collectant des données par Internet, par les applications mobiles, et même par **des capteurs placés dans le monde réel**, fournissent des données à LiveRamp. Certains d'entre eux utilisent les bases de données de LiveRamp, qui permettent aux entreprises « d'acheter et de vendre des données client précieuses ». D'autres fournissent des données afin que Acxiom et LiveRamp puissent reconnaître des individus et relier les informations enregistrées avec les profils numériques d'autres provenances. Mais le plus préoccupant, c'est sans doute le partenariat entre Acxiom et Crossix, une entreprise avec **des données détaillées sur la santé** de 250 millions de consommateurs américains. Crossix figure parmi les fournisseurs de données de LiveRamp.

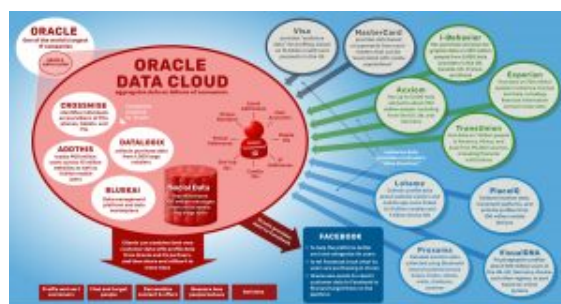
Quiconque enregistrant des données sur les consommateurs peut potentiellement être un fournisseur de données. »

Travis May, Directeur général de Acxiom-LiveRamp

Oracle, un géant des technologies de l'information pénètre le marché des données client

En faisant l'acquisition de plusieurs entreprises de données telles que Datalogix, BlueKai, AddThis et CrossWise, Oracle, un des premiers fournisseurs de logiciels d'entreprises et de bases de données dans le monde, est également récemment devenu **un des premiers courtiers en données clients**. Dans son « cloud », Oracle rassemble 3 milliards de profils utilisateurs issus de 15 millions de sites différents, les données d'un milliard d'utilisateurs mobiles, des milliards d'historiques d'achats dans des chaînes de supermarchés et 1500 détaillants, ainsi que 700 millions de messages par jour issus des réseaux sociaux, des blogs et des sites d'avis de consommateurs.

Oracle rassemble des données sur des milliards de consommateurs



Oracle et ses fournisseurs de données, ses partenaires et ses clients (en avril/mai 2017) - sources : voir le rapport

Oracle catalogue près de 100 fournisseurs de données dans son répertoire de données, parmi lesquels figurent Acxiom et des agences d'analyse de solvabilité telles que **Experian et TransUnion**, ainsi que des entreprises qui tracent les visites de sites Internet, l'utilisation d'applications mobiles et les déplacements, ou qui collectent des données à partir de questionnaires en ligne. **Visa et MasterCard** sont également référencés comme fournisseurs de données. En coopération avec ses partenaires, Oracle fournit plus de 30 000 catégories de données différentes qui peuvent être attribuées aux consommateurs. Réciproquement, l'entreprise partage des données avec Facebook et aide **Twitter** à calculer la solvabilité de ses utilisateurs.

Le **Graphe d'Identifiants** Oracle détermine et combine des profils utilisateur provenant de différentes entreprises. Il est le « trait d'union entre les interactions » à travers les différentes bases de données, services et appareils afin de « créer un profil client adressable » et « d'identifier partout les clients et les prospects ». D'autres entreprises peuvent envoyer à Oracle, des clés de correspondance construites à partir **d'adresses courriel, de numéros de téléphone, d'adresse postale ou d'autres identifiants**, Oracle les synchronisera ensuite à son « réseau d'identifiants utilisateurs et statistiques, connectés ensemble dans le Graphe d'Identifiants Oracle ». Bien que l'entreprise promette de n'utiliser que des identifiants utilisateurs anonymisés et des profils d'utilisateurs anonymisés, ceux-ci font tout de même référence à certains individus et peuvent être utilisés pour les reconnaître et les cibler dans de nombreux contextes de la vie.

Le plus souvent, les clients d'Oracle peuvent **télécharger dans le « cloud » d'Oracle leurs propres données concernant : leurs clients, les visites sur leur site ou les utilisateurs d'une application** ; ils peuvent les combiner avec des données issues de nombreuses autres entreprises, puis les transférer et les utiliser en temps réel sur des centaines d'autres plateformes de commerce et de publicité. Ils peuvent par exemple les utiliser pour **trouver et cibler** des personnes sur tous les appareils et plateformes, **personnaliser leurs interactions**, et le cas échéant **mesurer** la réaction des clients qui ont été personnellement ciblés.

La surveillance en temps réel des comportements quotidiens

Les plateformes en ligne, les fournisseurs de technologies publicitaires, les courtiers en données, et les négociants de toutes sortes d'industries peuvent maintenant surveiller, reconnaître et analyser des individus dans de nombreuses situations. Ils peuvent étudier ce qui intéresse les gens, ce qu'ils ont fait aujourd'hui, ce qu'ils vont sûrement faire demain, et leur valeur en tant que client.

Les données concernant les vies en ligne et hors ligne des

personnes

Une large spectre d'entreprises collecte des informations sur les personnes depuis des décennies. Avant l'existence d'Internet, les agences de crédit et les agences de marketing direct servaient de point d'intégration principal entre les données provenant de différentes sources. Une première étape importante dans **la surveillance systématique des consommateurs** s'est produite **dans les années 1990**, par la commercialisation de bases de données, les programmes de fidélité et l'analyse poussée de solvabilité. Après l'essor d'Internet et de la publicité en ligne **au début des années 2000**, et la montée des réseaux sociaux, des smartphones et de la publicité en ligne **à la fin des années 2000**, on voit maintenant **dans les années 2010** l'industrie des données clients s'intégrer avec le nouvel écosystème de pistage et de profilage numérique.

Cartographie de la collecte de données clients



Différents niveaux, domaines et sources de collecte de données clients par les entreprises

De longue date, les courtiers en données clients et d'autres entreprises acquièrent des informations sur **les abonnés à des journaux et à des magazines**, sur les membres de clubs de lecture et de ciné-clubs, sur les acheteurs de catalogues de vente par correspondance, sur les personnes réservant dans les agences de voyage, sur les participants à des séminaires et à des conférences, et sur les consommateurs qui remplissent les cartes de garantie pour leurs achats. La collecte de données d'achats grâce à des **programmes de fidélité** est, de ce point de vue, une pratique établie depuis longtemps.

En complément des données provenant directement des individus, sont utilisées,

par exemple les informations concernant **le type quartiers et d'immeubles** où résident les personnes afin de décrire, étiqueter, trier et catégoriser ces personnes. De même, les entreprises utilisent maintenant des profils de consommateurs s'appuyant sur les métadonnées concernant **le type de sites Internet fréquentés**, les vidéos regardées, les applications utilisées et les zones géographiques visitées. Au cours de ces dernières années, l'échelle et le niveau de détail des flux de données comportementales générées par toutes sortes d'activités du quotidien, telles que l'utilisation d'Internet, des réseaux sociaux et des équipements, ont rapidement augmenté.

Ce n'est pas un téléphone, c'est mon mouchard /pisteur/. New York Times, 2012

Un pistage et un profilage omniprésents

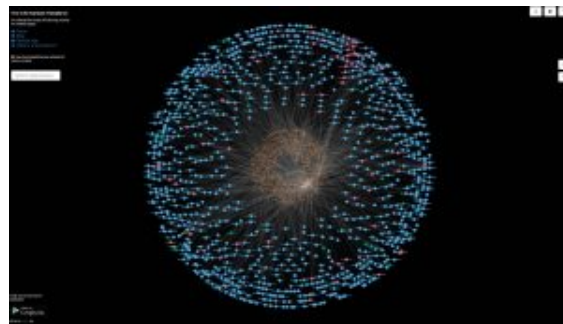
Une des principales raisons pour lesquelles le pistage et le profilage commerciaux sont devenus si généralisés c'est que quasiment tous les sites Internet, les fournisseurs d'applications mobiles, ainsi que de nombreux vendeurs d'équipements, partagent activement des données comportementales avec d'autres entreprises.

Il y a quelques années, la plupart des sites Internet ont commencé à inclure dans leur propre site **des services de pistage qui transmettent des données à des tiers**. Certains de ces services fournissent des fonctions visibles aux utilisateurs. Par exemple, lorsqu'un site Internet montre un bouton Facebook « j'aime » ou une vidéo YouTube encapsulée, des données utilisateur sont transmises à Facebook ou à Google. En revanche, de nombreux autres services ayant trait à la publicité en ligne demeurent cachés et, pour la plupart, ont pour seul objectif de collecter des données utilisateur. **Le type précis de données utilisateur** partagées par les éditeurs numériques et la façon dont les tierces parties utilisent ces données reste largement méconnus. Une partie de ces activités de pistage peut être analysée par n'importe qui ; par exemple en installant l'extension pour navigateur Lightbeam, il est possible de visualiser le réseau invisible des trackers des parties tierces.

Une étude récente a examiné un million de sites Internet différents et a trouvé plus de **80 000 services tiers** recevant des données concernant les visiteurs de ces sites. Environ 120 de ces services de pistage ont été trouvés sur plus de 10 000 sites, et six entreprises surveillent les utilisateurs sur plus de 100 000

sites, dont Google, Facebook, Twitter et BlueKai d'Oracle. Une étude sur 200 000 utilisateurs allemands visitant 21 millions de pages Internet a montré que les trackers tiers étaient présents **sur 95 %** des pages visitées. De même, la plupart des applications mobiles partagent des informations sur leurs utilisateurs avec d'autres entreprises. Une étude menée en 2015 sur les applications à la mode en Australie, en Allemagne et aux États-Unis a trouvé qu'**entre 85 et 95 % des applications gratuites**, et même **60 % des applications payantes** se connectaient à des tierces parties recueillant des données personnelles.

Une **carte interactive des services cachés de pistage tiers** sur les applications Android créée par des chercheurs européens et américains peut être explorée à l'adresse suivante : haystack.mobi/panopticon



Copie d'écran du ISCI Haystack Panopticon disponible sur haystack.mobi/panopticon, © mis à disposition gracieusement par ISCI, Université de Berkeley

En matière d'appareils, ce sont peut-être **les smartphones** qui actuellement contribuent le plus au recueil omniprésent données. L'information enregistrée par les téléphones portables fournit un aperçu détaillé de la personnalité et de la vie quotidienne d'un utilisateur. Puisque les consommateurs ont en général besoin d'un **compte Google, Apple ou Microsoft** pour les utiliser, une grande partie de l'information est déjà reliée à l'identifiant d'une des principales plateformes.

La vente de données utilisateurs ne se limite pas aux éditeurs de sites Internet et d'applications mobiles. Par exemple, l'entreprise d'intelligence commerciale SimilarWeb reçoit des données issues non seulement de centaines de milliers de sources de mesures directes depuis les sites et les applications, mais aussi des

logiciels de bureau et des extensions de navigateur. Au cours des dernières années, de nombreux autres appareils avec des capteurs et des connexions réseau ont intégré la vie de tous les jours, **cela va des liseuses électroniques et autres accessoires connectés aux téléphones intelligents, compteurs, thermostats, détecteurs de fumée, imprimantes, réfrigérateurs, brosses à dents, jouets et voitures**. À l'instar des smartphones, ces appareils donnent aux entreprises un accès sans précédent au comportement des consommateurs dans divers contextes de leur vie.

Publicité programmatique et technologie marketing

La plus grande partie de la publicité numérique prend aujourd'hui la forme d'enchères en temps réel hautement automatisées entre les éditeurs et les publicitaires ; on appelle cela la **publicité programmatique**. Lorsqu'une personne se rend sur un site Internet, les données utilisateur sont envoyées à une kyrielle de services tiers, qui cherchent ensuite à reconnaître la personne et extraire l'information disponible sur le profil. Les publicitaires souhaitant livrer une publicité à cet individu, en particulier du fait de certains attributs ou comportements, placent une enchère. En quelques millisecondes, le publicitaire le plus offrant gagne et place la pub. Les publicitaires peuvent de la même façon **enchérir sur les profils utilisateurs** et le placement de publicités au sein des applications mobiles.

Néanmoins, ce processus ne se déroule pas, la plupart du temps, entre **les éditeurs et les publicitaires**. L'écosystème est constitué d'une pléthore de toutes sortes de données différentes et de fournisseurs de technologies en interaction les uns avec les autres, parmi lesquels des réseaux publicitaires, des marchés publicitaires, des plateformes côté vente et des plateformes côté achat. Certains se spécialisent dans le pistage et la publicité suivant les résultats de recherche, dans la publicité généraliste sur Internet, dans la pub sur mobile, dans les pubs vidéos, dans les pubs sur les réseaux sociaux, ou dans les pubs au sein des jeux. D'autres se concentrent sur l'approvisionnement en données, en analyse ou en services de personnalisation.

Pour tracer le portrait des utilisateurs d'Internet et d'applications mobiles, toutes les parties impliquées ont développé des méthodes sophistiquées pour accumuler, regrouper et relier les informations provenant de différentes entreprises afin **de suivre les individus dans tous les aspects de leur vie**. Nombre d'entre elles

recueillent et utilisent des profils numériques sur des centaines de millions de consommateurs, leurs navigateurs Internet et leurs appareils.

De nombreux secteurs rejoignent l'économie de pistage

Au cours de ces dernières années, des entreprises dans plusieurs secteurs ont commencé à partager et à utiliser à très grande échelle des données concernant leurs utilisateurs et clients.

La plupart des **détaillants** vendent des formes agrégées de **données sur les habitudes d'achat** auprès des entreprises d'études de marchés et des courtiers en données. Par exemple, l'entreprise de données IRI accède aux données de plus de 85 000 magasins ('alimentation, grande distribution, médicaments, d'alcool et d'animaux de compagnie, magasin à prix unique et magasin de proximité). Nielsen déclare recueillir les informations concernant les ventes de 900 000 magasins dans le monde dans plus de 100 pays. L'enseigne de grande distribution britannique **Tesco** sous-traite son programme de fidélité et ses activités en matière de données auprès d'une filiale, Dunnhumby, dont le slogan est « transformer les données consommateur en régal pour le consommateur ». Lorsque Dunnhumby a fait l'acquisition de l'entreprise technologique de publicité allemande Sociomantic, il a été annoncé que Dunnhumby « conjuguerait ses connaissances étendues au sujet sur les préférences d'achat de 400 millions de consommateurs » avec les « données en temps réel de plus de 700 millions de consommateurs en ligne » de Sociomantic afin de personnaliser et d'évaluer les publicités.

Cartographie de l'écosystème du pistage et du profilage commercial



Aujourd'hui de nombreux industriels dans divers secteurs ont rejoint l'écosystème de pistage et

de profilage numérique, aux cotés des grandes plateformes en ligne et des professionnels de l'analyse des données clients.

De grands **groupes médiatiques** sont aussi fortement intégrés dans l'écosystème de pistage et de profilage numérique actuel. Par exemple, Time Inc. a fait l'acquisition d'Adelphic, une importante société de pistage et de technologies publicitaires multi-support, mais aussi de Viant, une entreprise qui déclare avoir accès à plus de 1,2 milliard d'utilisateurs enregistrés. La plateforme de streaming **Spotify** est un exemple célèbre d'éditeur numérique qui vend les données de ses utilisateurs. Depuis 2016, la société partage avec le département données du géant du marketing WPP des informations à propos de ce que les utilisateurs écoutent, sur leur humeur ainsi que sur leur comportement et leur activité en termes de playlist. WPP a maintenant accès « aux préférences et comportements musicaux des 100 millions d'utilisateurs de Spotify ».

De nombreuses grandes **entreprises de télécom et de fournisseurs d'accès Internet** ont fait l'acquisition d'entreprises de technologies publicitaires et de données. Par exemple, Millennial Media, une filiale d'**AOL-Verizon**, est une plateforme de publicité mobile qui collecte les données de plus de 65 000 applications de différents développeurs, et prétend avoir accès à environ 1 milliard d'utilisateurs actifs distincts dans le monde. **Singtel**, l'entreprise de télécoms basée à Singapour, a acheté Turn, une plateforme de technologies publicitaires qui donne accès aux distributeurs à 4,3 milliards d'appareils pouvant être ciblés et d'identifiants de navigateurs et à 90 000 attributs démographiques, comportementaux et psychologiques.

Comme les compagnies aériennes, les hôtels, les commerces de détail et les entreprises de beaucoup d'autres secteurs, le secteur des **services financiers** a commencé à agréger et utiliser des données clients supplémentaires grâce à des programmes de fidélité dans les années 80 et 90. Les entreprises dont la clientèle cible est proche et complémentaires partagent depuis longtemps certaines de leurs données clients entre elles, un processus souvent géré par des intermédiaires. Aujourd'hui, l'un de ces intermédiaires est Cardlytics, une entreprise qui gère des programmes de fidélité pour plus de 1 500 institutions financières, telles que **Bank of America** et MasterCard. Cardlytics s'engage

auprès des institutions financières à « générer des nouvelles sources de revenus en exploitant le pouvoir de [leurs] historiques d'achat ». L'entreprise travaille aussi en partenariat avec LiveRamp, la filiale d'Acxiom qui combine les données en ligne et hors ligne des consommateurs.

Pour **MasterCard**, la vente de produits et de services issus de l'analyse de données pourrait même devenir son **cœur de métier**, sachant que la production d'informations, dont la vente de données, représentent une part considérable et croissante de ses revenus. **Google** a récemment déclaré qu'il capture environ 70 % des transactions par carte de crédit aux États-Unis via « partenariats tiers » afin de tracer les achats, mais n'a pas révélé ses sources.

Ce sont vos données. Vous avez le droit de les contrôler, de les partager et de les utiliser comme bon vous semble.

C'est ainsi que le courtier en données Lotame s'adresse sur son site Internet à ses entreprises clientes en 2016.

Relier, faire correspondre et combiner des profils numériques

Jusqu'à récemment, les publicitaires, sur Facebook, Google ou d'autres réseaux de publicité en ligne, ne pouvaient cibler les individus qu'en analysant leur comportement en ligne. Mais depuis quelques années, grâce aux moyens offerts par les entreprises de données, les profils numériques issus de différentes plateformes, de différentes bases de données clients et du monde de la publicité en ligne peuvent désormais être associés et combinés entre eux.

Connecter les identités en ligne et hors ligne

Cela a commencé en 2012, quand Facebook a permis aux entreprises de télécharger leurs propres **listes d'adresses de courriel et de numéros de téléphone** sur la plateforme. Bien que les adresses et numéros de téléphone soient convertis en pseudonyme, Facebook est en mesure de relier directement ces données client provenant d'entreprises tierces avec ses propres comptes utilisateur. Cela permet par exemple aux entreprises de **trouver et de cibler très précisément sur Facebook** les personnes dont elles possèdent les adresses de courriel ou les numéros de téléphone. De la même façon, il leur est

éventuellement possible d'exclure certaines personnes du ciblage de façon sélective, ou de déléguer à la plateforme le repérage des personnes qui ont des caractéristiques, centre d'intérêts, et comportements communs.

C'est une **fonctionnalité puissante**, peut-être plus qu'il n'y paraît au premier abord. Elle permet en effet aux entreprises d'associer systématiquement leurs données client avec les données Facebook. Mieux encore, d'autres publicitaires et marchands de données peuvent également synchroniser leurs bases avec celles de la plateforme et en exploiter les ressources, ce qui équivaut à fournir une sorte de **télécommande en temps réel pour manipuler l'univers des données Facebook**. Les entreprises peuvent maintenant capturer en temps réel des données comportementales extrêmement précises comme un clic de souris sur un site, le glissement d'un doigt sur une application mobile ou un achat en magasin, et demander à Facebook de trouver et de cibler aussitôt les personnes qui viennent de se livrer à ces activités. Google et Twitter ont mis en place des fonctionnalités similaires en 2015.

Les plateformes de gestion de données

De nos jours, la plupart des entreprises de technologie publicitaire croisent en continu plusieurs sources de codage relatives aux individus. Les **plateformes de gestion de données** permettent aux entreprises de tous les domaines **d'associer et de relier leurs propres données clients**, comprenant des informations en temps réel sur les achats, les sites web consultés, les applications utilisées et les réponses aux courriels, avec des profils numériques fournis par une multitude de fournisseurs tiers de données. Les données associées peuvent alors être analysées, triées et classées, puis utilisées pour envoyer un message donné à des personnes précises via des réseaux ou des appareils particuliers. Une entreprise peut, par exemple, cibler un groupe de clients existants ayant visité une page particulière sur son site ; ils sont alors **perçus comme pouvant devenir de bons clients**, bénéficiant alors de contenus personnalisés ou d'une réduction, que ce soit sur Facebook, sur une appli mobile ou sur le site même de l'entreprise.

L'émergence des plateformes de gestion de données **marque un tournant** dans le développement d'un envahissant pistage des comportements d'achat. Avec leur aide, les entreprises dans tous les domaines et partout dans le monde peuvent **très facilement associer et relier** les données qu'elles ont collectées depuis des

années sur leurs clients et leurs prospects avec les milliards de profils collectés dans le monde numérique. Les principales entreprises faisant tourner ces plateformes sont : Oracle, Adobe, Salesforce (KruX), Wunderman (KBM Group/Zipline), Neustar, Lotame et Cxense.

Nous vous afficherons des publicités basées sur votre identité, mais cela ne veut pas dire que vous serez identifiable.

Erin Egan, Directeur de la protection de la vie privée chez Facebook, 2012

Identifier les gens et relier les profils numériques

Pour surveiller et suivre les gens dans les différentes situations de leur vie, pour leur associer des profils et toujours les reconnaître comme un seul et même individu, les entreprises amassent une grande variété de types de données qui, en quelque sorte, les identifient.

Parce qu'il est ambigu, le nom d'une personne a toujours été un mauvais identifiant pour un recueil de données. **L'adresse postale**, par contre, a longtemps été et est encore, une indication clé qui permet d'associer et de relier des données de différentes origines sur les consommateurs et leur famille. Dans le monde numérique, les identifiants les plus pertinents pour relier les profils et les comportements sur les différentes bases de données, plateformes et appareils sont : **l'adresse de courriel, le numéro de téléphone**, et le code propre à chaque smartphone ou autre appareil.

Les **identifiants de compte utilisateur** sur les immenses plateformes comme Google, Facebook, Apple et Microsoft jouent aussi un rôle important dans le suivi des gens sur Internet. Google, Apple, Microsoft et Roku attribuent un « **identifiant publicitaire** » aux individus, qui est maintenant largement utilisé pour faire correspondre et relier les données d'appareils tels que les smartphones avec les autres informations issues du monde numérique. Verizon utilise son propre identifiant pour pister les utilisateurs sur les sites web et les appareils. Certaines grandes entreprises de données comme **Acxiom, Experian et Oracle** disposent, au niveau mondial, d'un identifiant unique par personne qu'elles utilisent pour relier des dizaines d'années de données clients avec le monde numérique. Ces **identifiants d'entreprise** sont constitués le plus souvent de deux identifiants ou plus qui sont attachés à différents aspects de la vie en ligne

et hors ligne d'une personne et qui peuvent être d'une certaine façon reliés l'un à l'autre.

Des Identifiants utilisés pour pister les gens sur les sites web, les appareils et les lieux de vie



Comment les entreprises identifient les consommateurs et les relient à des informations de profils - sources : voir le rapport

Les entreprises de pistage utilisent également des identifiants plus ou moins temporaires, comme les cookies qui sont attachés aux utilisateurs surfant sur le web. Depuis que les utilisateurs peuvent ne pas autoriser ou supprimer les cookies dans leur navigateur, elles ont développé des méthodes sophistiquées permettant de calculer une empreinte numérique unique basée sur diverses caractéristiques du navigateur et de l'ordinateur d'une personne. De la même manière, les entreprises amassent les empreintes sur les appareils tels que les smartphones. Les cookies et les empreintes numériques sont **continuellement synchronisés** entre les différents services de pistage et ensuite reliés à des identifiants plus permanents.

D'autres entreprises fournissent des services de pistage multi-appareils qui utilisent le *machine learning* (voir Wikipédia) pour analyser de grandes quantités de données. Par exemple, **Tapad**, qui a été acheté par le géant des télécoms norvégiens Telenor, analyse les données de deux milliards d'appareils dans le monde et utilise des modèles basés sur les comportements et les relations pour trouver la probabilité qu'un ordinateur, une tablette, un téléphone ou un autre appareil appartienne à la même personne.

Un profilage « anonyme » ?

Les entreprises de données suppriment les noms dans leurs profils détaillés et utilisent des fonctions de hachage (voir Wikipedia) pour convertir les adresses de courriel et les numéros de téléphone en **code alphanumérique** comme "e907c95ef289". Cela leur permet de déclarer sur leur site web et dans leur politique de confidentialité qu'elles recueillent, partagent et utilisent uniquement des données clients « anonymisées » ou « dé-identifiées ».

Néanmoins, comme la plupart des entreprises utilisent les mêmes process déterministes pour calculer ces codes alphanumériques, on devrait les considérer comme des pseudonymes qui sont en fait bien plus pratiques que les noms réels pour identifier les clients dans le monde numérique. Même si une entreprise partage des profils contenant uniquement des adresses de courriels ou des numéros de téléphones chiffrés, **une personne peut toujours être reconnue** dès qu'elle utilise un autre service lié avec la même adresse de courriel ou le même numéro de téléphone. De cette façon, bien que chaque service de pistage impliqué ne connaissent qu'une partie des informations du profil d'une personne, les entreprises peuvent **suivre et interagir avec les gens** au niveau individuel via les services, les plateformes et les appareils.

Si une entreprise peut vous suivre et interagir avec vous dans le monde numérique - et cela inclut potentiellement votre téléphone mobile ou votre télé - alors son affirmation que vous êtes anonyme n'a aucun sens, en particulier quand des entreprises ajoutent de temps à autre des informations hors-ligne aux données en ligne et masquent simplement le nom et l'adresse pour rendre le tout « anonyme ».

Joseph Turow, spécialiste du marketing et de la vie privée dans son livre « The Daily You », 2011

Gérer les clients et les comportements : personnalisation et évaluation

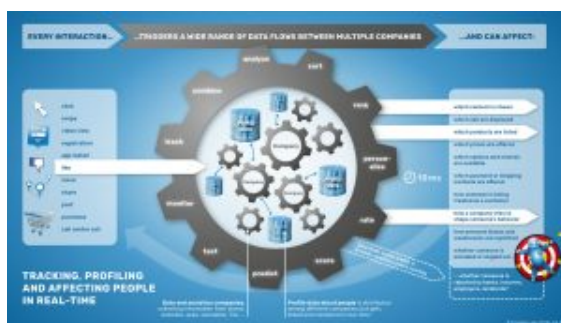
S'appuyant sur les méthodes sophistiquées d'interconnexion et de combinaison de données entre différents services, les entreprises de tous les secteurs d'activité peuvent utiliser les **flux de données comportementales actuellement**

omniprésents afin de surveiller et d'analyser une large gamme d'activités et de comportements de consommateurs pouvant être pertinents vis-à-vis de leurs intérêts commerciaux.

Avec l'aide des vendeurs de données, les entreprises tentent d'entrer en contact avec les clients tout au long de leurs parcours autant de fois que possible, à travers les achats en ligne ou en boutique, le publipostage, les pubs télé et les appels des centres d'appels. Elles tentent **d'enregistrer et de mesurer chaque interaction** avec un consommateur, y compris sur les sites Internet, plateformes et appareils qu'ils ne contrôlent pas eux-mêmes. Elles peuvent recueillir en continu une abondance de données concernant leurs clients et d'autres personnes, les améliorer avec des informations provenant de tiers, et utiliser les profils améliorés au sein de l'écosystème de commercialisation et de technologie publicitaire. À l'heure actuelle, les plateformes de gestion des données clients permettent **la définition de jeux complexes de règles** qui régissent la façon de réagir automatiquement à certains critères tels que des activités ou des personnes données ou une combinaison des deux.

Par conséquent, les **individus ne savent jamais** si leur comportement a déclenché une réaction de l'un de ces réseaux de pistage et de profilage constamment mis à jour, interconnectés et opaques, ni, le cas échéant, comment cela influence les options qui leur sont proposées à travers les canaux de communication et dans les situations de vie.

Tracer, profiler et influencer les individus en temps réel



Chaque interaction enclenche un large éventail de flux de données entre de nombreuses entreprises.

Personnalisation en série

Les flux de données échangés entre les publicitaires en ligne, les courtiers en données, et les autres entreprises ne sont pas seulement utilisés pour diffuser de la publicité ciblée sur les sites web ou les applis mobiles. Ils sont de plus en plus utilisés pour **personnaliser les contenus, les options et les choix** offerts aux consommateurs sur le site d'une entreprise par exemple. Les entreprises de technologie des données, comme par exemple Optimizely, peuvent aider à personnaliser un site web spécialement pour les personnes qui le visitent pour la première fois, en s'appuyant sur les profils numériques de ces visiteurs fournis par Oracle.

Les boutiques en ligne, par exemple, personnalisent l'accueil des visiteurs : **quels produits** seront mis en évidence, **quelles promotions** seront proposées, et même le prix et des produits ou des services peuvent être différents selon la personne qui visite le site. Les services de détection de la fraude évaluent les utilisateurs en temps réel et décident **quels moyens de paiement et de transport** peuvent être proposés.

Les entreprises développent des technologies pour calculer et évaluer en continu le potentiel de valeur à long terme d'un client en s'appuyant sur son historique de navigation, de recherche et de localisation, mais aussi sur son usage des applis, sur les produits achetés et sur ses amis sur les réseaux sociaux. Chaque clic, chaque glissement de doigt, chaque *Like*, chaque partage est susceptible d'influencer **la manière dont une personne est traitée en tant que client**, combien de temps elle va attendre avant que la *hotline* ne lui réponde, ou si elle sera complètement exclue des relances et des services marketing.

L'Internet des riches n'est pas le même que celui des pauvres.

Michael Fertik, fondateur de reputation.com, 2013

Trois types de **plateformes technologiques** jouent un rôle important dans cette sorte de personnalisation instantanée. Premièrement, les entreprises utilisent des systèmes de gestion de la relation client pour gérer leurs données sur les clients et les prospects. Deuxièmement, elles utilisent des plateformes de gestion de données pour connecter leurs propres données à l'écosystème de publicité numérique et obtiennent ainsi des informations supplémentaires sur le profil de

leurs clients. Troisièmement, elles peuvent utiliser des **plateformes de marketing prédictif** qui les aident à produire le bon message pour la bonne personne au bon moment, calculant comment convaincre quelqu'un en exploitant ses faiblesses et ses préjugés.

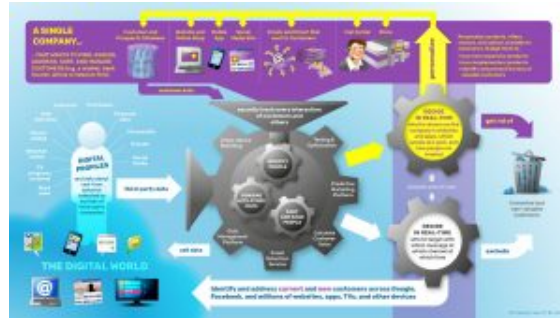
Par exemple, l'entreprise de données **RocketFuel** promet à ses clients de « leur apporter des milliers de milliards de signaux numériques ou non pour créer des profils individuels et pour fournir aux consommateurs une expérience personnalisée, toujours actualisée et toujours pertinente » s'appuyant sur les 2,7 milliards de profils uniques de son dépôt de données. Selon RocketFuel, il s'agit « de noter chaque signal selon sa propension à influencer le consommateur ».

La plateforme de marketing prédictif **TellApart**, qui appartient à **Twitter**, associe une valeur à chaque couple client/produit acheté, une « synthèse entre la probabilité d'achat, l'importance de la commande et la valeur à long terme », s'appuyant sur « des centaines de signaux en ligne et en magasin sur un consommateur anonyme unique ». En conséquence, TellApart regroupe automatiquement du contenu tel que « l'image du produit, les logos, les offres et toute autre métadonnée » pour construire des publicités, des courriels, des sites web et des offres personnalisées.

Tarifs personnalisés et campagnes électorales

Des méthodes identiques peuvent être utilisées pour personnaliser les tarifs dans les boutiques en ligne, par exemple, en prédisant le niveau d'achat d'un client à long terme ou le montant qu'il sera probablement prêt à payer un peu plus tard. Des preuves sérieuses suggèrent que les boutiques en ligne affichent déjà des **tarifs différents selon les consommateurs**, ou même des prix différents pour le même produit, en s'appuyant sur leur comportement et leurs caractéristiques. Un champ d'action similaire est la personnalisation lors des campagnes électorales. Le ciblage des électeurs avec des messages personnalisés, adaptés à leur personnalité, et à leurs opinions politiques sur des problèmes donnés a fait monter les débats sur une possible **manipulation politique**.

Utiliser les données, les analyser et les personnaliser pour gérer les consommateurs



Actuellement, dans tous les domaines, les entreprises peuvent mobiliser les réseaux de suivi et de profilage pour trouver, évaluer, contacter, trier et gérer les consommateurs

Tests et expériences sur les personnes

La personnalisation s'appuyant sur de riches informations de profil et sur du suivi invasif en temps réel est devenue **un outil puissant pour influencer le comportement du consommateur** quand il visite une page web, clique sur une pub, s'inscrit à un service, s'abonne à une newsletter, télécharge une application ou achète un produit.

Pour améliorer encore cela, les entreprises ont commencé à faire des expériences en continu sur les individus. Elles procèdent à **des tests en faisant varier** les fonctionnalités, le design des sites web, l'interface utilisateur, les titres, les boutons, les images ou mêmes les tarifs et les remises, surveillent et mesurent avec soin comment les différents groupes d'utilisateurs interagissent avec ces modifications. De cette façon, les entreprises **optimisent sans arrêt leur capacité à encourager les personnes** à agir comme elles veulent qu'elles agissent.

Les organes de presse, y compris à grand tirage comme le Washington Post, utilisent **différentes versions des titres de leurs articles** pour voir laquelle est la plus performante. Optimizely, un des principaux fournisseurs de technologies pour ce genre de tests, propose à ses clients la capacité de « faire des tests sur l'ensemble de l'expérience client sur n'importe quel canal, n'importe quel appareil, et n'importe quelle application ». Expérimenter sur des usagers qui

l'ignorent **est devenu la nouvelle norme.**

En 2014, **Facebook** a déclaré faire tourner « plus d'un millier d'expérimentations chaque jour » afin « d'optimiser des résultats précis » ou pour « affiner des décisions de design sur le long terme ». En 2010 et 2012, la plateforme a mené des expérimentations sur des millions d'utilisateurs et montré qu'en manipulant l'interface utilisateur, les fonctionnalités et le contenu affiché, Facebook **pouvait augmenter significativement le taux de participation électorale** d'un groupe de personnes. Leur célèbre expérimentation sur l'humeur des internautes, portant sur 700 000 individus, consistait à manipuler secrètement la quantité de messages émotionnellement positifs ou négatifs présents dans les fils d'actualité des utilisateurs : il s'avéra que cela avait **un impact sur le nombre de messages positifs ou négatifs** que les utilisateurs postaient ensuite eux-mêmes.

Suite à la critique massive de Facebook par le public concernant cette expérience, la plateforme de rendez-vous **OkCupid** a publié un article de blog provocateur défendant de telles pratiques, déclarant que « nous faisons des expériences sur les êtres humains » et « c'est ce que font tous les autres ». OkCupid a décrit une expérimentation dans laquelle a été manipulé le pourcentage de « compatibilité » montré à des paires d'utilisateurs. Quand on affichait un taux de 90 % entre deux utilisateurs qui en fait étaient peu compatibles, les utilisateurs échangeaient nettement plus de messages entre eux. OkCupid a déclaré que quand elle « dit aux gens » qu'ils « vont bien ensemble », alors ils « agissent comme si c'était le cas ».

Toutes ces expériences qui posent de vraies questions éthiques montrent le **pouvoir de la personnalisation basée sur les données pour influencer sur les comportements.**

Dans les mailles du filet : vie quotidienne, données commerciales et analyse du risque

Les données concernant les comportements des personnes, les liens sociaux, et les moments les plus intimes sont de plus en plus utilisées dans des contextes ou à des fins complètement différents de ceux dans lesquels elles ont été enregistrées. Notamment, elles sont de plus en plus utilisées pour prendre des

décisions automatisées au sujet d'individus dans des domaines clés de la vie tels que la finance, l'assurance et les soins médicaux.

Données relatives aux risques pour le marketing et la gestion client

Les agences d'évaluation de la solvabilité, ainsi que d'autres acteurs clés de l'évaluation du risque, principalement dans des domaines tels que la **vérification des identités, la prévention des fraudes, les soins médicaux et l'assurance** fournissent également des solutions commerciales. De plus, la plupart des courtiers en données s'échangent divers types d'informations sensibles, par exemple des informations concernant la situation financière d'un individu, et ce à des fins commerciales. **L'utilisation de l'évaluation de solvabilité à des fins de marketing** afin soit de cibler soit d'exclure des ensembles vulnérables de la population a évolué pour devenir des produits qui associent le marketing et la gestion du risque.

L'agence d'évaluation de la solvabilité **TransUnion** fournit, par exemple, un **produit d'aide à la décision piloté par les données** à destination des commerces de détail et des services financiers qui leur permet « de mettre en œuvre des stratégies de marketing et de gestion du risque sur mesure pour atteindre les objectifs en termes de clients, canaux de vente et résultats commerciaux », il inclut des données de crédit et promet « un aperçu inédit du comportement, des préférences et des risques du consommateur. » Les entreprises peuvent alors laisser leurs clients « choisir parmi une gamme complète d'offres sur mesure, répondant à leurs besoins, leurs préférences et leurs profils de risque » et « évaluer leurs clients sur divers produits et canaux de vente et leur présenter uniquement la ou les offres les plus pertinentes pour eux et les plus rentables » pour l'entreprise. De même, **Experian** fournit un produit qui associe « crédit à la consommation et informations commerciales, fourni avec plaisir par Experian. »

En matière de surveillance, il n'est pas question de connaître vos secrets, mais de gérer des populations, de gérer des personnes.

Katarzyna Szymielewicz, Vice-Présidente EDRI, 2015

Vérification des identités en ligne et détection de la fraude

Outre la machine de surveillance en temps réel qui a été développée au travers de la publicité en ligne, d'autres formes de pistage et de profilage généralisées ont émergé dans les domaines de **l'analyse de risque, de la détection de fraudes et de la cybersécurité**.

De nos jours, les services de détection de fraude en ligne utilisent des technologies hautement intrusives afin **d'évaluer des milliards de transactions numériques**. Ils recueillent d'énormes quantités d'informations concernant les appareils, les individus et les comportements. Les fournisseurs habituels dans l'évaluation de solvabilité, la vérification d'identité, et la prévention des fraudes ont commencé à surveiller et à évaluer la façon dont les personnes surfent sur le web et utilisent leurs appareils mobiles. En outre, ils ont entrepris **de relier les données comportementales en ligne avec l'énorme quantité d'information hors-connexion** qu'ils recueillent depuis des dizaines d'années.

Avec l'émergence de services passant par l'intermédiaire d'objets technologiques, la vérification de l'identité des consommateurs et la prévention de la fraude sont devenues de plus en plus importantes et de plus en plus contraignantes, notamment au vu de la cybercriminalité et de la fraude automatisée. Dans un même temps, les systèmes actuels d'analyse du risque ont agrégé **des bases de données gigantesques contenant des informations sensibles** sur des pans entiers de population. Nombre de ces systèmes répondent à un grand nombre de cas d'utilisation, parmi lesquels la preuve d'identité pour les services financiers, l'évaluation des réclamations aux compagnies d'assurance et des demandes d'indemnités, de l'analyse des transactions financières et l'évaluation de milliards de transactions en ligne.

De tels systèmes d'analyse du risque peuvent décider si une requête ou une transaction est **acceptée ou rejetée** ou décider des options de livraison disponibles pour une personne lors d'une transaction en ligne. Des services marchands de vérification d'identité et d'analyse de la fraude sont également employés dans des domaines tels que **les forces de l'ordre et la sécurité nationale**. La frontière entre les applications commerciales de l'analyse de l'identité et de la fraude et celles utilisées par les agences gouvernementales de renseignement est de plus en plus floue.

Lorsque des individus sont ciblés par des systèmes aussi opaques, ils peuvent être **signalés comme étant suspects et nécessitant un traitement particulier ou une enquête**, ou bien ils peuvent être rejetés sans plus d'explication. Ils peuvent recevoir un courriel, un appel téléphonique, une notification, un message d'erreur, ou bien le système peut tout simplement ne pas indiquer une option, sans que l'utilisateur ne connaisse son existence pour d'autres. Des évaluations erronées peuvent se propager d'un système à l'autre. Il est souvent **difficile, voire impossible de faire recours** contre ces évaluations négatives qui excluent ou rejettent, notamment à cause de la difficulté de s'opposer à quelque chose dont on ne connaît pas l'existence.

Exemples de détection de fraude en ligne et de service d'analyse des risques

L'entreprise de cybersécurité **ThreatMetrix** traite les données concernant 1,4 milliard de « comptes utilisateur uniques » sur des « milliers de sites dans le monde. » Son Digital Identity Network (Réseau d'Identité Numérique) enregistre des « millions d'opérations faites par des consommateurs chaque jour, notamment des connexions, des paiements et des créations de nouveaux comptes », et cartographie les « associations en constante évolution entre les individus et leurs appareils, leurs positions, leurs identifiants et leurs comportements » à des fins de vérification des identités et de prévention des fraudes. L'entreprise collabore avec Equifax et TransUnion. Parmi ses clients se trouvent Netflix, Visa et des entreprises dans des secteurs tels que le jeu vidéo, les services gouvernementaux et la santé.

De façon analogue, l'entreprise de données **ID Analytics**, qui a récemment été achetée par **Symantec**, exploite un Réseau d'Identifiants fait de « 100 millions de nouveaux éléments d'identité quotidiens issus des principales organisations interprofessionnelles. ». L'entreprise agrège des données concernant 300 millions de consommateurs, sur les prêts à haut risque, les achats en ligne et les demandes de carte de crédit ou de téléphone portable. Son Indice d'Identité, ID Score, prend en compte les appareils numériques ainsi que les noms, les numéros de sécurité sociale et les adresses postales et courriel.

Trustev, une entreprise en ligne de détection de la fraude dont le siège se situe en Irlande et qui a été rachetée par l'agence d'évaluation de la solvabilité **TransUnion** en 2015, juge des transactions en ligne pour des clients dans les

secteurs des services financiers, du gouvernement, de la santé et de l'assurance en s'appuyant sur l'analyse des comportements numériques, les identités et les appareils tels que les téléphones, les tablettes, les ordinateurs portables, les consoles de jeux, les télévisions et même les réfrigérateurs. L'entreprise propose aux entreprises clientes la possibilité d'analyser **la façon dont les visiteurs cliquent et interagissent avec les sites Internets et les applications**. Elle utilise une large gamme de données pour évaluer les utilisateurs, y compris les numéros de téléphone, les adresses courriel et postale, les empreintes de navigateur et d'appareil, les vérifications de la solvabilité, les historiques d'achats sur l'ensemble des vendeurs, les adresses IP, les opérateurs mobiles et la géolocalisation des téléphones. Afin d'aider à « accepter les transactions futures », chaque appareil se voit attribuer une empreinte digitale d'appareil unique. Trustev propose aussi une technologie de marquage d'empreinte digitale sociale qui analyse le contenu des réseaux sociaux, notamment une « analyse de la liste d'amis » et « l'identification des schémas ». TransUnion a intégré la technologie Trustev dans ses propres solutions identifiantes et anti-fraude.

Selon son site Internet, Trustev utilise une large gamme de données pour évaluer les personnes



Capture d'écran du site Internet de Trustev, 2 juin 2016

De façon similaire, l'agence d'évaluation de la solvabilité Equifax affirme qu'elle possède des données concernant près de 1 milliard d'appareils et peut affirmer « l'endroit où se situe en fait un appareil et s'il est associé à d'autres appareils utilisés dans des fraudes connues ». En associant ces données avec « des milliards d'identités et d'événements de crédit pour trouver les activités douteuses » dans tous les secteurs, et en utilisant des informations concernant **la situation**

d'emploi et les liens entre les ménages, les familles et les partenaires, Equifax prétend être capable « de distinguer les appareils ainsi que les individus ».

Je ne suis pas un robot

Le produit **reCaptcha** de Google fournit en fait un service similaire, du moins en partie. Il est incorporé dans des millions de sites Internets et aide les fournisseurs de sites Internets à décider si un visiteur est un être humain ou non. Jusqu'à récemment, les utilisateurs devaient résoudre diverses sortes de défis rapides tels que le déchiffrement de lettres dans une image, la sélection d'images dans une grille, ou simplement en cochant la case « Je ne suis pas un robot ». En 2017, **Google** a présenté une version invisible de reCaptcha, en expliquant qu'à partir de maintenant, les utilisateurs humains pourront passer « **sans aucune interaction utilisateur**, contrairement aux utilisateurs douteux et aux robots ». L'entreprise ne révèle pas le type de données et de comportements utilisateurs utilisés pour reconnaître les humains. Des analyses laissent penser que Google, outre les adresses IP, les empreintes de navigateur, la façon dont l'utilisateur frappe au clavier, déplace la souris ou utilise l'écran tactile « avant, pendant et après » une interaction reCaptcha, utilise plusieurs témoins Google. **On ne sait pas exactement** si les individus sans compte utilisateur sont désavantagés, si Google est capable d'identifier des individus particuliers plutôt que des « humains » génériques, ou si Google utilise les données enregistrées par reCaptcha à d'autres fins que la détection de robots.

Le pistage numérique à des fins publicitaires et de détection de la fraude ?

Les flux omniprésents de données comportementales enregistrées pour la publicité en ligne s'écoulent vers les systèmes de détection de la fraude. Par exemple, **la plateforme de données commerciales** Segment propose à ses clients des moyens faciles d'envoyer des données concernant leurs clients, leur site Internet et les utilisateurs mobiles à une kyrielle de services de technologies commerciales, ainsi qu'à des **entreprises de détection de fraude**. Castle est l'une d'entre-elles et utilise « les données comportementales des consommateurs pour prédire les utilisateurs qui présentent vraisemblablement un risque en matière de sécurité ou de fraude ». Une autre entreprise, Smyte, aide à « prévenir les arnaques, les messages indésirables, le harcèlement et les fraudes par carte

de crédit ».

La grande agence d'analyse de la solvabilité **Experian** propose un service de pistage multi-appareils qui fournit de la reconnaissance universelle d'appareils, sur mobile, Internet et les applications **pour le marketing numérique**. L'entreprise s'engage à concilier et à associer les « identifiants numériques existants » de leurs clients, y compris des « témoins, identifiants d'appareil, adresses IP et d'autres encore », fournissant ainsi aux commerciaux un « lien omniprésent, cohérent et permanent sur tous les canaux ».

La technologie d'identification d'appareils provient de **41st parameter** (le 41e paramètre), une entreprise de détection de la fraude rachetée par Experian en 2013. En s'appuyant sur la technologie développée par 41st parameter, Experian propose aussi une solution d'intelligence d'appareil pour **la détection de la fraude** au cours des paiements en ligne. Cette solution qui « crée un identifiant fiable pour l'appareil et recueille des données appareil abondantes » « identifie en quelques millisecondes chaque appareil à chaque visite » et « fournit une visibilité jamais atteinte de l'individu réalisant le paiement ». On ne sait pas exactement si Experian utilise les mêmes données pour ses services d'identification d'appareils pour détecter la fraude que pour le marketing.

Cartographie de l'écosystème du pistage et du profilage commercial

Au cours des dernières années, les pratiques déjà existantes de surveillance commerciale ont rapidement muté en un large éventail d'acteurs du secteur privé qui surveillent en permanence des populations entières. Certains des acteurs de l'écosystème actuel de pistage et de profilage, tels que **les grandes plateformes** et d'autres **entreprises avec un grand nombre de clients**, tiennent une position unique en matière d'étendue et de niveau de détail de leurs profils de consommateurs. Néanmoins, les données utilisées pour prendre des décisions concernant les individus sur de nombreux sujets ne sont généralement pas centralisées en un lieu, mais plutôt assemblées en temps réel à partir de plusieurs sources selon les besoins.

Un large éventail **d'entreprises de données et de services d'analyse** en marketing, en gestion client et en analyse du risque recueillent, analysent, partagent et échangent de façon uniforme des données client et les associent avec

des informations supplémentaires issues de milliers d'autres entreprises. Tandis que l'industrie des données et des services d'analyse fournissent les moyens pour déployer ces puissantes technologies, **les entreprises dans de nombreuses industries** contribuent à augmenter la quantité et le niveau de détail des données collectées ainsi que la capacité à les utiliser.

Cartographie de l'écosystème du pistage et du profilage commercial numérique



En plus des grandes plateformes en ligne et de l'industrie des données et des services d'analyse des consommateurs, des entreprises dans de nombreux secteurs ont rejoint les écosystèmes de pistage et de profilage numérique généralisé.

Google et Facebook, ainsi que d'autres grandes plateformes telles que Apple, Microsoft, Amazon et Alibaba ont un accès sans précédent à des données concernant les vies de milliards de personnes. Bien qu'ils aient des modèles commerciaux différents et jouent par conséquent des rôles différents dans l'industrie des données personnelles, ils ont le pouvoir de dicter dans une large mesure les paramètres de base des marchés numériques globaux. Les grandes plateformes limitent principalement la façon dont les autres entreprises peuvent obtenir leurs données. Ainsi, ils les obligent à utiliser les **données utilisateur de la plateforme dans leur propre écosystème** et recueillent des données au-delà de la portée de la plateforme.

Bien que **les grandes multinationales** de différents secteurs ayant des

interactions fréquentes avec des centaines de millions de consommateurs soient en quelque sorte dans une situation semblable, elles ne font pas qu'acheter des données clients recueillies par d'autres, elles en fournissent aussi. Bien que certaines parties des secteurs des services financiers et des télécoms ainsi que des **domaines sociétaux critiques** tels que la santé, l'éducation et l'emploi soient soumis à une réglementation plus stricte dans la plupart des juridictions, un large éventail d'entreprises a commencé à utiliser ou fournissent des données aux réseaux actuels de surveillance commerciale.

Les détaillants et d'autres entreprises qui vendent des produits et services aux consommateurs vendent pour la plupart les données concernant les achats de leurs clients. **Les conglomérats médiatiques et les éditeurs numériques** vendent des données au sujet de leur public qui sont ensuite utilisées par des entreprises dans la plupart des autres secteurs. **Les fournisseurs de télécoms et d'accès haut débit** ont entrepris de suivre leurs clients sur Internet. Les grandes groupes de distribution, de médias et de télécoms ont acheté ou achètent **des entreprises de données, de pistage et de technologie publicitaire**. Avec le rachat de NBC Universal par Comcast et le rachat probable de Time Warner par AT&T, les grands groupes de télécoms aux États-Unis sont aussi en train de devenir des éditeurs gigantesques, créant par là même des portefeuilles puissants de contenu, de données et de capacité de pistage. Avec l'acquisition de AOL et de Yahoo, Verizon aussi est devenu une « plateforme ».

Les institutions financières ont longtemps utilisé des données sur les consommateurs pour la gestion du risque, notamment dans l'évaluation de la solvabilité et la détection de fraude, ainsi que pour le marketing, l'acquisition et la rétention de clientèle. Elles complètent leurs propres données avec des données externes issues d'agences d'évaluation de la solvabilité, de courtiers en données et d'entreprises de données commerciales. **PayPal**, l'entreprise de paiements en ligne la plus connue, partage des informations personnelles avec plus de 600 tiers, parmi lesquels d'autres fournisseurs de paiements, des agences d'évaluation de la solvabilité, des entreprises de vérification de l'identité et de détection de la fraude, ainsi qu'avec les acteurs les plus développés au sein de l'écosystème de pistage numérique. Tandis que les réseaux de cartes de crédit et les banques ont partagé des informations financières sur leurs clients avec les fournisseurs de données de risque depuis des dizaines d'années, ils ont maintenant commencé à **vendre des données sur les transactions** à des fins

publicitaires.

Une myriade d'entreprises, grandes ou petites, **fournissant des sites Internets, des applications mobiles, des jeux et d'autres solutions** sont étroitement liées à l'écosystème de données commerciales. Elles utilisent des services qui leur permettent de facilement transmettre à des services tiers des données concernant leurs utilisateurs. Pour nombre d'entre elles, la vente de flux de données comportementales concernant leurs utilisateurs constitue un élément clé de leur *business model*. De façon encore plus inquiétante, les entreprises qui fournissent des services tels que les **enregistreurs d'activité physique** intègrent des services qui transmettent les données utilisateurs à des tierces parties.

L'envahissante machine de surveillance en temps réel qui a été développée pour **la publicité en ligne** est en train de s'étendre vers d'autres domaines dont la politique, la tarification, la notation des crédits et la gestion des risques. Partout dans le monde, **les assureurs** commencent à proposer à leurs clients des offres incluant du suivi en temps réel de leur comportement : comment ils conduisent, quelles sont leurs activités santé ou leurs achats alimentaires et quand ils se rendent au club de gym. Des nouveaux venus dans **l'analyse assurantielle et les technologies financières** prévoient les risques de santé d'un individu en s'appuyant sur les données de consommation, mais évaluent aussi la solvabilité à partir de données de comportement via les appels téléphoniques ou les recherches sur Internet.

Les courtiers en données sur les consommateurs, les entreprises de gestion de clientèle et les agences de publicité comme Acxiom, Epsilon, Merkle ou Wunderman/WPP jouent un rôle prépondérant en rassemblant et reliant les données entre les plateformes, les multinationales et le monde de la technologie publicitaire. **Les agences d'évaluation de crédit** comme Experian qui fournissent de nombreux services dans des domaines très sensibles comme l'évaluation de crédit, la vérification d'identité et la détection de la fraude jouent également un rôle prépondérant dans l'actuel envahissant écosystème de la commercialisation des données.

Des entreprises particulièrement importantes qui fournissent des données, des analyses et des solutions logicielles sont également appelées « plateforme ». **Oracle**, un fournisseur important de logiciel de base de données est, ces dernières années, devenu un courtier en données de consommation. **Salesforce**,

le leader sur le marché de la gestion de la relation client qui gère les bases de données commerciales de millions de clients qui ont chacun de nombreux clients, a récemment acquis Krux, une grande entreprise de données, connectant et combinant des données venant de l'ensemble du monde numérique. L'entreprise de logiciels **Adobe** joue également un rôle important dans le domaine des technologies de profilage et de publicité.

En plus, les principales grandes entreprises **du conseil, de l'analyse et du logiciel commercial**, comme IBM, Informatica, SAS, FICO, Accenture, Capgemini, Deloitte et McKinsey et même des **entreprises spécialisées dans le renseignement et la défense** comme Palantir, jouent également un rôle significatif dans la gestion et l'analyse des données personnelles, de la gestion de la relation client à celle de l'identité, du marketing à l'analyse de risque pour les assureurs, les banques et les gouvernements.

Vers une société du contrôle social numérique généralisé ?

Ce rapport montre qu'aujourd'hui, les réseaux entre plateformes en ligne, fournisseurs de technologies publicitaires, courtiers en données, et autres peuvent **suivre, reconnaître et analyser des individus dans de nombreuses situations de la vie courante**. Les informations relatives aux comportements et aux caractéristiques d'un individu sont reliées entre elles, assemblées, et utilisées en temps réel par des entreprises, des bases de données, des plateformes, des appareils et des services. Des acteurs uniquement motivés par des buts économiques ont fait naître un environnement de données dans lequel les individus sont constamment **sondés et évalués, catégorisés et regroupés, notés et classés, numérotés et comptés, inclus ou exclus**, et finalement traités de façon différente.

Ces dernières années, plusieurs **évolutions importantes** ont donné de nouvelles capacités sans précédent à la surveillance omniprésente par les entreprises. Cela comprend l'augmentation des médias sociaux et des appareils en réseau, le pistage et la mise en relation en temps réel de flux de données comportementales, le rapprochement des données en ligne et hors ligne, et la consolidation des données commerciales et de gestion des risques. L'envahissant pistage et profilage numériques, mélangé à la personnalisation et aux tests, ne sont pas seulement utilisés pour surveiller, mais aussi pour influencer systématiquement **le comportement des gens**. Quand les entreprises utilisent les données sur les

situations du quotidien pour prendre des décisions parfois triviales, parfois conséquente sur les gens, cela peut conduire à **des discriminations**, et renforcer voire aggraver des inégalités existantes.

Malgré leur omniprésence, seul le haut de l'iceberg des données et des activités de profilage est visible pour les particuliers. La plupart d'entre elles restent opaques et à peine compréhensible par la majorité des gens. Dans le même temps, les gens ont **de moins en moins de solutions pour résister au pouvoir de cet écosystème de données** ; quitter le pistage et le profilage envahissant, est devenu synonyme de quitter la vie moderne. Bien que les responsables des entreprises affirment que la vie privée est morte (tout en prenant soin de préserver leur propre vie privée), Mark Andrejevic suggère que les gens perçoivent en fait l'asymétrie du pouvoir dans le monde numérique actuel, mais se sentent « frustrés par un sentiment d'impuissance face à une collecte et à une exploitation de données de plus en plus sophistiquées et exhaustives. »

Au regard de cela, ce rapport se concentre sur le fonctionnement interne et les pratiques en vigueur dans l'actuelle industrie des données personnelles. Bien que l'image soit devenue plus nette, de larges portions du système restent encore dans le noir. **Renforcer la transparence sur le traitement des données par les entreprises** reste un prérequis indispensable pour résoudre le problème de l'asymétrie entre les entreprises de données et les individus. Avec un peu de chance, les résultats de ce rapport encourageront des travaux ultérieurs de la part de journalistes, d'universitaires, et d'autres personnes concernés par les libertés civiles, la protection des données et celle des consommateurs ; et dans l'idéal des travaux des législateurs et des entreprises elles-mêmes.

En 1999, Lawrence Lessig, avait bien prédit que, laissé à lui-même, le cyberspace, deviendrait un parfait outil de contrôle façonné principalement par la « main invisible » du marché. Il avait dit qu'il était possible de « construire, concevoir, ou programmer le cyberspace pour protéger les valeurs que nous croyons fondamentales, ou alors de construire, concevoir, ou programmer le cyberspace pour permettre à toutes ces valeurs de disparaître. » De nos jours, la deuxième option est presque devenue réalité au vu des milliards de dollars investis dans le capital-risque pour financer des modèles économiques s'appuyant sur une exploitation massive et sans scrupule des données. L'insuffisance de régulation sur la vie privée aux USA et l'absence de son application en Europe ont réellement gêné l'émergence **d'autres modèles d'innovation numérique**, qui

seraient fait de pratiques, de technologies, de modèles économiques qui protègent la liberté, la démocratie, la justice sociale et la dignité humaine.

À un niveau plus global, la législation sur la protection des données ne pourra pas, à elle seule, atténuer les conséquences qu'un monde « conduit par les données » a sur les individus et la société que ce soit aux USA ou en Europe. Bien que le **consentement et le choix** soient des principes cruciaux pour résoudre les problèmes les plus urgents liés à la collecte massive de données, ils peuvent également mener à une illusion de volontarisme. En plus d'instruments de régulation supplémentaires sur la non-discrimination, la protection du consommateur, les règles de concurrence, il faudra en général **un effort collectif important** pour donner une vision positive d'une future société de l'information. Sans quoi, on pourrait se retrouver bientôt dans une société avec un envahissant contrôle social numérique, dans la laquelle la vie privée deviendrait, si elle existe encore, un luxe pour les riches. Tous les éléments en sont déjà en place.

Lectures pour approfondir le sujet

— L'article ci-dessus en format .PDF (376,2 Ko)

-> framablog.org-Comment les entreprises surveillent notre quotidien

— Un essai plus exhaustif sur les questions abordées par la publication ci-dessus ainsi que des références et des sources peuvent être trouvés dans le rapport complet, disponible au téléchargement en PDF.

— Le rapport de 2016 « Les réseaux du contrôle » par Wolfie Christl et Sarah Spiekermann sur lequel le présent rapport est largement fondé est disponible au téléchargement en PDF ainsi qu'en format papier.

La production de ce rapport, matériaux web et illustrations a été soutenue par Open Society Foundations.

Bibliographie

Christl, W. (2017, juin). Corporate surveillance in everyday life. Cracked Labs.

Christl, W., & Spiekermann, S. (2016). *Networks of Control, a Report on Corporate Surveillance, Digital Tracking, Big Data & Privacy* (p. 14-20).

Consulté à l'adresse

https://www.privacylab.at/wp-content/uploads/2016/09/Christl-Networks_K_o.pdf

Epp, C., Lippold, M., & Mandryk, R. L. (2011). Identifying emotional states using keystroke dynamics (p. 715). ACM Press.

<https://doi.org/10.1145/1978942.1979046>

Kosinski, M., Stillwell, D., & Graepel, T. (2013). Private traits and attributes are predictable from digital records of human behavior. *Proceedings of the National Academy of Sciences*, 110(15), 5802-5805.

<https://doi.org/10.1073/pnas.1218772110>

Turow, J. (s. d.). *Daily You* | Yale University Press. Consulté 25 septembre 2017, à l'adresse

<https://yalebooks.yale.edu/book/9780300188011/daily-you>