



HAL
open science

À l'ombre des rencontres : la fabrique des données personnelles sur OkCupid

Camille Girard-Chanudet

► **To cite this version:**

Camille Girard-Chanudet. À l'ombre des rencontres : la fabrique des données personnelles sur OkCupid. Les Cahiers du numérique, 2021, 17 (1-2), pp.45-69. 10.3166/LCN.2021.005 . hal-04648520

HAL Id: hal-04648520

<https://hal.science/hal-04648520v1>

Submitted on 15 Jul 2024

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

À L'OMBRE DES RENCONTRES

La fabrique des données personnelles sur OkCupid

Camille Girard-Chanudet. 2021. « À l'ombre des rencontres : la fabrique des données personnelles sur OkCupid », Les Cahiers du Numérique, vol.17, n°1-2, pp 45-69.

À partir de l'étude de cas du site de rencontres OkCupid, cet article rend compte des processus de création des données personnelles, à la charnière entre plateforme et utilisateurs. L'analyse de l'environnement économique, graphique et technique de création des données, conçu en *back-office* par la plateforme, permet de mieux comprendre les rapports inégaux que les internautes entretiennent individuellement et collectivement avec ces objets informationnels. Cet article vise également à mettre en évidence la multiplicité des sens attribués à ces objets immatériels et mouvants, dont le rôle d'intermédiaire est essentiel pour le fonctionnement de services numériques articulant des ensembles hétérogènes d'acteurs.

Mots-clés : Données personnelles ; Vie privée ; Sites de rencontre ; Plateformes

Based on the case study of a dating app – OkCupid –, this article focuses on the processes of personal data construction that take place at the meeting point between platforms and their users. The analysis of the economic, graphic and technical environment in which data creation take place – designed in the platform's back-office – allows to better understand the uneven relation users individually and collectively develop with these informational objects. This article also aims at highlighting the multiple meanings that these immaterial and moving objects bear, as well as the essential role they play as intermediaries in the functioning of digital services articulating heterogeneous groups of actors.

Keywords : Personal data ; Privacy ; Dating apps ; Platform

1. Introduction

Les données personnelles numériques occupent une place centrale et croissante dans le fonctionnement de services connectés, portés par des plateformes de types diversifiés. Dans les domaines des transports, de la mode, de la finance ou encore des rencontres amoureuses, une grande partie des activités en ligne repose sur l'exploitation de ces objets aux multiples facettes : les modèles économiques (notamment publicitaires) des plateformes dépendent de l'activation de leur valeur commerciale (Crain, 2018; Rochelandet, 2010), tandis que leur potentiel analytique oriente les stratégies et évolutions d'un grand nombre d'entreprises. L'exploitation algorithmique des données par des systèmes dits d'« intelligence artificielle » conduit enfin à structurer des communautés d'utilisateurs autour de services numériques personnalisés (Cardon, 2019; Just & Latzer, 2017; Reigeluth, 2018).

Dans un univers numérique longtemps marqué par les valeurs libertaires d'horizontalité et de libre circulation des informations (Loveluck, 2015), la prise de conscience de l'existence de ces processus de captation et d'utilisation d'informations personnelles par des acteurs quasi-hégémoniques a fait surgir d'importantes craintes quant aux potentialités de surveillance et de contrôle numérique (voir par exemple Sadin, 2010; Zuboff, 2015). Deux ensembles de recherches se sont rapidement structurés autour de cette problématique : les travaux du courant des « *surveillance studies* », d'une part, analysent les impacts politiques de l'émergence de tels instruments de contrôle (Lyon et al., 2012). D'autre part, un ensemble de publications s'est organisé autour de la question des usages, en cherchant à résoudre le « paradoxe de la vie privée » (Barnes, 2006), par lequel les individus s'asserviraient volontairement aux grandes plateformes en leur transmettant des informations à caractère personnel les concernant (Acquisti, 2011; Nissenbaum, 2011; Turow et al., 2015). Ces travaux ont en commun leur appréhension des « données personnelles » comme des objets tangibles aux frontières clairement définies.

Cet article propose de déplacer cette perspective, en remontant aux racines de l'objet « données personnelles » et en questionnant sa construction. Il s'agira de faire l'effort de ne pas considérer les données personnelles comme des réalités hors-sol et figées que les individus pourraient choisir ou non d'échanger et que les plateformes exploiteraient, mais de s'intéresser aux mécanismes sociotechniques de façonnage de ces objets de connaissance. Ces processus se mettent en place au point de rencontre entre individus et plateformes, et nécessitent l'articulation d'acteurs pluriels au sein d'environnements spécifiques (Denis, 2018). Ils contribuent ainsi à la structuration de communautés au sein

desquelles le travail relationnel, intermédié technologiquement, se trouve « distribué » (Licoppe, 2007). C'est pourquoi nous ne parlerons pas de collecte ou de partage de données qui seraient déjà là, mais adopterons au contraire une approche processuelle s'incarnant dans les termes de création et de construction des données, ou encore de « donnéification ».

Nous entendrons ainsi par « donnéification » – tiré de l'anglais « datafication » (Van Dijck, 2014) –, les processus de traduction d'éléments constitutifs de l'identité des individus en objets informationnels standardisés, enregistrables, quantifiables et traitables automatiquement : les données. Ces opérations se déroulent dans un cadre organisationnel, économique, graphique et technique particulier, que nous nommerons « espace de donnéification ».

Une telle approche conduira à mettre en évidence la nature particulière des données personnelles numériques, qui diffèrent fortement des objets d'étude traditionnels des sociologies s'intéressant aux relations entre humains et artefacts. Il s'agit d'objets à la matérialité incertaine, dont la création et l'existence se déroulent dans l'ombre d'outils techniques premièrement dédiés à d'autres activités (en particulier les applications et sites, eux-mêmes soumis à leurs supports physiques, ordinateurs ou téléphones). Les internautes ne font en effet pas usage des données personnelles, mais plutôt de services numériques sur lesquels sont greffées en *back-office* les opérations de donnéification, dont les résultats impactent en retour les expériences vécues individuellement et collectivement sur les interfaces.

L'émergence des données personnelles gagne de ce fait à être étudiée dans le cadre de la relation entre plateforme et utilisateurs. L'exercice consistera à articuler ces deux plans : une approche en termes d'innovation et de techniques (Hughes, 1983; MacKenzie, 1990) permettra d'une part de cerner la construction de l'environnement économique, graphique et technique encadrant la donnéification côté plateformes. C'est en prenant en compte ce contexte que seront analysés, d'autre part, les enjeux d'appropriation et de réception liés aux données personnelles, dans une perspective axée cette fois vers les usages (Akrich, 1990; Jouët, 2000).

En cherchant à tenir ensemble ces deux niveaux, cet article cherchera à éclairer par le prisme des données personnelles les interactions entre plateformes numériques et individus et communautés – dont ont déjà été documentées, entre autres, les dimensions professionnelles (Chaves Ferreira et al., 2018; Rosenblat, 2018), algorithmiques (Beuscart et al., 2019; Hallinan & Striphos, 2016; Möller et al., 2018) ou encore de sociabilité (Bastard et al., 2017). L'orientation vers les données personnelles permettra en particulier de mettre l'accent sur la

multiplicité et l'enchevêtrement des niveaux d'activité dans lesquels les plateformes et leurs utilisateurs sont engagés.

Cette recherche s'ancre dans une étude de cas, celui de la plateforme OkCupid. Acteur majeur du secteur de la rencontre créé en 2004, OkCupid a pour particularité de revendiquer un appariement par affinité de ses utilisateurs, sur la base des réponses fournies par ceux-ci à un grand nombre de questions personnelles. Nous avons mené une enquête de terrain sur cette plateforme entre 2017 et 2019, en combinant différentes méthodes d'enquête visant à rendre compte à la fois des logiques structurantes de la plateforme, et des pratiques et appréciations des utilisateurs au sein de l'espace de donnification.

Côté plateforme, le design de l'interface a été analysé par le biais d'un processus de rétro-ingénierie orienté vers l'identification de ses caractéristiques architecturales (organisation et accessibilité des différentes fonctionnalités) et graphiques (choix, forme et disposition des éléments). Cette observation a été complétée par l'étude de divers corpus documentaires donnant à voir les modalités organisationnelles, financières et techniques de la création de données sur OkCupid (rapports trimestriels d'activité destinés aux actionnaires de la maison-mère IAC Corp, contrats enregistrés à la Security and Exchange Commission, différentes versions successives de la politique de confidentialité). Enfin, les valeurs structurantes de la plateforme ont été abordées par l'intermédiaire d'articles du blog d'OkCupid et du livre *Datachysm* publié par l'un de ses fondateurs.

Côté utilisateurs, une enquête par entretiens et par questionnaire a été réalisée auprès de membres recrutés directement sur la plateforme. 45 entretiens de nature directive ont été conduits en ligne sur son espace de *chat* et 11 entretiens semi-dirigés d'une durée moyenne d'une heure et demie ont été réalisés hors ligne. Enfin, un questionnaire concernant leurs pratiques, connaissances et perceptions a été rempli par 250 membres (échantillon constitué paritairement d'hommes et de femmes, âgés de 19 à 45 ans et membres d'OkCupid depuis une durée allant de quelques jours à 5 ans).

Sur la base de cette enquête, cet article vise dans un premier temps à dessiner les contours du cadre de création des données au sein duquel évoluent les utilisateurs d'OkCupid. Il s'agira d'exposer les caractéristiques de l'environnement économique (réseau de plateformes intégré à l'« économie des données »), structurel (design et architecture de l'interface) et technique (mécanismes de captation des informations personnelles) qui sous-tend la production des données personnelles en *back-office*. Ceci nous permettra de mettre en évidence le rôle prépondérant joué par la plateforme dans la structuration de

l'espace de donnéification, ainsi que dans la production et la circulation des données.

Ce panorama servira de base pour identifier, dans un second temps, la place occupée par les utilisateurs dans les processus de donnéification. En nous intéressant à la perception qu'ont les membres d'OkCupid de ces mécanismes, nous montrerons les limites structurelles à l'engagement conscient des utilisateurs dans la donnéification, et soulignerons l'influence de différents contextes, temps d'usage et appartenances sociales sur leur appréhension de ces processus, ayant pour conséquence des variations importantes d'exposition à la donnéification entre individus.

2. Le cadre de création des données personnelles

Les données personnelles numériques sont des objets informationnels, construits à partir de la rencontre entre internautes – dévoilant des fragments de leur identité en utilisant des services numériques – et plateformes – captant et convertissant ces informations en éléments standardisés et exploitables. Comprendre l'émergence des données suppose de s'intéresser à l'environnement spécifique de cette rencontre, qui se structure en plusieurs niveaux interdépendants. Il s'agit en premier lieu d'un cadre économique et organisationnel, celui de l'écosystème de plateformes sous-tendant ce qui a été nommé l'« économie des données » (Rochelandet, 2010), dans lequel s'inscrit l'assemblage d'entreprises propriétaires d'OkCupid. C'est au cœur de cet écosystème qu'est conçue la partie visible de la plateforme pour les utilisateurs : l'interface sur laquelle ils évoluent. Second niveau du cadre de création des données, cet espace est caractérisé par des choix d'architecture et de design influençant les comportements et perceptions des membres. Le troisième niveau, greffé sur l'interface, est constitué par les outils techniques de donnéification (*cookies, web beacons...*) opérant la transformation concrète d'éléments constitutifs de l'identité des individus en objets numériques calculables. C'est à ces trois niveaux du cadre de création de données personnelles sur OkCupid, et à l'impact de leurs caractéristiques sur la donnéification, que nous allons nous intéresser dans cette première partie. Cette étude permettra de mettre en évidence le rôle décisif joué par la plateforme dans la configuration de l'environnement dans lequel les données sont créées, et son impact sur les possibilités d'engagement des individus dans ces processus.

2.1. OkCupid, tête de proue d'un assemblage d'entreprises intégré à l'économie des données

Le service de rencontres OkCupid est piloté, comme nombre de plateformes numériques, par un agrégat d'entreprises, prenant part à des niveaux distincts à la conception et à la mise en œuvre des processus de donnification. Si la gestion courante de la plateforme (maintenance de l'interface, service client, gestion des systèmes d'information) est assurée par l'entreprise Humour Rainbow Inc, les orientations générales sont décidées au niveau de son entreprise-mère, MatchGroup, elle-même filiale d'Inter Active Corporation (IAC). Le groupe Match est propriétaire de la majorité des grandes plateformes de rencontre (Tinder, Meetic, Pairs...), ce qui lui permet à la fois de segmenter librement le marché de la rencontre, et d'équilibrer la fréquentation de ses différentes plateformes en se réservant le droit de basculer des profils des uns vers les autres, comme stipulé dans leurs conditions d'utilisation (version 2018). IAC est quant à elle, malgré sa discrétion dans le paysage médiatique, une entreprise étasunienne de taille détenant plus de 150 marques dans des domaines variés. L'opacité de ces structures est proportionnelle au pouvoir décisionnaire et économique dont elles disposent : si une lecture consciencieuse des conditions d'utilisation peut renseigner les utilisateurs sur l'existence d'Humour Rainbow Inc, le rôle d'IAC n'est en revanche jamais mentionné dans les documents communiqués aux membres.

Les entreprises gestionnaires d'OkCupid font partie d'un écosystème économique plus large centré sur les données personnelles. En effet, si les données sont utilisées en interne pour la personnalisation algorithmique des services proposés, leur revente est également une source majeure de revenus pour les entreprises participant à leur création. Cette activité est au cœur du secteur en forte croissance de la publicité ciblée numérique, nourrie par la circulation de données personnelles au sein de réseaux de plateformes intermédiées par des *data hubs*.

A titre d'illustration, dans le cas d'OkCupid, l'entreprise IAC a conclu en 2016 un accord commercial avec le *data hub* Ad Words, branche publicitaire de Google, lui réservant jusqu'en 2020 l'ensemble des espaces publicitaires des plateformes qu'elle possède, et permettant le ciblage personnalisé des publicités présentées par le biais de l'intégration aux interfaces de cookie-traceurs liés à Google. Cet accord, impliquant le transfert de données personnelles des membres vers le *data hub* de Google, fournit à IAC Corp presque un quart de ses revenus annuels, et met OkCupid en lien avec les millions d'autres plateformes connectées au réseau AdWords¹.

¹ Sources : Security and Exchange Commission, IAC shareholder letters

La construction des données prend donc place au sein d'un espace dense peuplé de structures économiques multiples (plateformes, *holdings*, *hubs*). L'émergence des données, et leur « passivation », c'est-à-dire leur conversion en biens suffisamment stabilisés et qualifiés pour être échangés (Callon, 2013), représente un enjeu stratégique et économique majeur pour les acteurs évoluant dans cet écosystème. C'est à cette lumière qu'on peut comprendre le travail de cadrage de l'environnement graphique et architectural à partir duquel les données sont constituées.

2.2. L'interface, masque souriant favorisant l'exposition à la donnification

De ce large écosystème, les utilisateurs ne perçoivent que la partie émergée : l'interface web ou mobile. C'est sur elle qu'ils s'inscrivent, parcourent les différents profils, initient des contacts. Les caractéristiques spécifiques de l'interface d'OkCupid contribuent à invisibiliser les tâches techniques de donnification effectuées en *back-office* (Denis, 2018; Star & Strauss, 1999). Elles orientent l'attention des internautes et conditionnent leur appréciation de la plateforme, par le biais de stratégies de personnification, d'injonction et de ludification.

OkCupid cherche ainsi à apparaître aux yeux de ses utilisateurs sous une forme humanisée, bienveillante et joviale. Tout nouveau membre est accueilli par un message automatique d'Alice, une employée d'OkCupid disposant d'un profil sur le site. Par la suite, l'utilisation du service est ponctuée par la présence de petits personnages guidant de façon humoristique le parcours l'internaute. Cette personnification de la plateforme contribue à générer chez les utilisateurs une impression de connivence avec elle, comme l'exprime Louise, une utilisatrice de 30 ans voyant en OkCupid « un compagnon sympa [lui faisant] rencontrer des gens sympas ». Cette identité factice nourrit les représentations des internautes et se substitue de ce fait dans les imaginaires à la réalité administrative et financière de la plateforme. C'est en lien avec ces choix de conception de l'interface que l'on peut interpréter le fait que plusieurs membres, à l'instar de Jean (31 ans, communiquant), considèrent que leur vie privée est davantage protégée sur OkCupid que d'autres sites de rencontre pourtant soumis aux mêmes conditions de donnification car appartenant à la même maison-mère :

T'as toujours une sorte de confidentialité fausse, parce que tu sais que les données peuvent se revendre à des tiers. Encore qu'OkCupid est pas trop trop pourri sur ce point-là, beaucoup moins que les autres sites de rencontre.

Ce « compagnon sympa » n'est pas neutre pour les membres de la plateforme : c'est par son biais qu'OkCupid leur prodigue des « conseils » favorisant la création de données, dont le caractère injonctif est masqué par la légèreté du ton. S'il est indispensable de ne fournir qu'un nombre restreint d'informations pour pouvoir utiliser OkCupid (genre, orientation sexuelle, âge, mail, ville, photo, quelques mots de description, réponse à 15 questions), la forme de l'interface encourage une plus grande exposition des utilisateurs. Celle-ci rythme en effet leur parcours de légères incitations : « Réponds à plus de questions », « Écris un peu à propos de toi », « Décris tes loisirs », « Ajoute ton compte Instagram », « Envoie-lui un message tout de suite »... Réponds, écris, décris, ajoute, envoie... Ces « injonctions à l'action » (Tubaro & Casilli, 2017) suggèrent aux internautes des utilisations intensives et extraverties de la plateforme qui contribuent de ce fait à l'élargissement du champ de création des données personnelles.

L'élargissement de cette base est également le fruit de la ludification de l'interface, dont les fonctionnalités sont graphiquement et architecturalement conçues pour stimuler l'attention et l'engagement des utilisateurs par le plaisir et le défi (Deterding, 2012). L'un des fondateurs d'OkCupid explique ainsi la mise en œuvre de cette stratégie (traduction de l'auteure) (Rudder, 2014) :

Je peux vous le dire de l'intérieur : les entreprises conçoivent leurs produits de façon à susciter l'engagement. Sur OkCupid, vous voyez le nombre de messages reçus, de visites, vous pouvez mesurer vos possibilités. On sait que ces chiffres stimulent votre intérêt, surtout quand ils gonflent. Sans une dose d'excitation, les sites ou les applications ont l'air morts et tout le monde s'en va.

L'une des fonctionnalités les plus ludifiées d'OkCupid est celle consistant à répondre à des séquences de questions à choix multiples appelés « questions de match ». Elle place les utilisateurs dans des boucles d'engagement constituées d'enchaînements de tâches courtes et simples (répondre en cochant des cases) dont la réalisation est gratifiante : répondre permet en particulier de débloquent diverses options (pourcentage de match, analyse de la personnalité...). Cette fonctionnalité sert de base à la création de stocks très importants de données personnelles, puisqu'elle consiste de fait à faire spécifier leurs goûts, pratiques et préférences aux individus puis à enregistrer et traiter ces informations. Toutefois, lorsque ceux-ci utilisent cette fonctionnalité, ils ne le font pas dans l'objectif de collaborer à cette donnification, mais y recourent plutôt, pour la plupart, à des fins de diversion, parce que répondre à des questions de match est « rigolo » et permet de « faire passer le temps », comme l'explique Clément (34 ans, ingénieur) :

- Question : Quand tu réponds à des questions de match, pourquoi tu le fais, qu'est-ce qui te motive ?
- Clément : Bah, ce qui me motive c'est l'ennui... (rire). Je suis assez occupé, mais de temps en temps faut que je fasse des pauses (...). Ça me permet de faire un truc assez court, et j'ai pas d'impératif de temps dessus ou de trucs comme ça.
- Question : Mais tu le fais pas pour améliorer le pourcentage de match par exemple ?
- Clément : Non. Je pense pas que les gens regardent les questions. Enfin... Y'a tellement de questions qu'honnêtement...

L'organisation par la plateforme de l'espace visuel et architectural oriente donc l'appréciation, l'attention et les comportements des internautes, favorisant leur exposition à la donnification sans pour autant les impliquer consciemment dans ces processus.

C'est derrière cette façade que sont mis en place les moyens techniques de donnification permettant de convertir les déclarations et pratiques des internautes en objets standardisés.

2.3. La douceur des mécanismes techniques de donnification

De façon générale, les moyens techniques de donnification mis en œuvre sur le web sont rarement visibles. En dehors des champs à compléter fréquemment présents lors de l'inscription à un service ou des questions posées comme sur OkCupid, la création de données se fait principalement par le biais de technologies discrètes ciblant les traces laissées par les individus sur les interfaces (Rouvroy & Berns, 2010).

Parmi les 32 types de techniques de traçage identifiés par une équipe de recherche en science informatique (Bujlow et al., 2017), la plateforme OkCupid en met en œuvre au moins deux, qu'elle mentionne dans sa politique de confidentialité : les *cookies* et les *web beacons*.

Les *cookies* sont de petits fichiers texte générés par un serveur et stockés, par le biais du navigateur, sur l'ordinateur de l'internaute lorsque celui-ci visite un site. Ils contiennent diverses informations telles que les logins, préférences ou historique de navigation, et sont employés pour « se souvenir » de chaque utilisateur chaque fois qu'il accède au service. Les *cookies* appartiennent généralement au domaine sur lequel l'internaute se trouve (*first-party cookies*), mais ils peuvent également communiquer des informations à d'autres entités, en particulier lorsque le site intègre de la publicité ciblée (*third-party cookies*). Sur

OkCupid, au moins 15 sites tiers, parmi lesquels le réseau Google Ad Words, Facebook et des entreprises de marketing digital (Hexagon Analytics), collectent des informations à propos des membres avant même que ceux-ci se connectent².

La seconde technique, les *web beacons*, sont de petites images invisibles, mesurant généralement 1x1 pixels, intégrées aux interfaces des sites internet ou aux emails. Lorsque l'utilisateur ouvre une page contenant un *web beacon*, son navigateur sollicite l'image au serveur correspondant, et transmet dans le même temps un certain nombre d'informations (adresse IP, type de navigateur, historique de navigation...). En retour, un cookie est placé sur le dispositif de l'utilisateur, permettant de pouvoir tracer son parcours sur le web au fil des différents sites présentant la même petite image.

Ces deux méthodes de donnéification reposent sur des protocoles HTTP ou Java, utilisés comme canaux de communication entre clients et serveurs. Ces langages ne sont pas conçus pour être compris ni même visualisés par l'internaute moyen. Au contraire, les techniques de traçage numérique ont vocation à être transparentes, et opérer en arrière-plan des activités conduites par les internautes sans les perturber.

En ce sens, ces méthodes relèvent de la « surveillance douce » identifiée par Gary Marx : parce qu'elle se joue en grande partie à un niveau imperceptible pour les utilisateurs, la création de données ne génère pas de sentiment de contrainte chez les internautes (Marx, 2006). Cette configuration pose question en termes de capacité d'action des individus face à la création des données personnelles, puisqu'elle limite leur connaissance de ce processus à des abstractions théoriques détachées des expériences vécues ou ressenties sur l'interface ; or il est difficile de participer – ou de résister – à quelque chose que l'on ne voit ni ne perçoit.

À l'immatérialité des mécanismes de donnéification s'ajoute le fait que leur produit soit, lui aussi, difficile à appréhender pour les utilisateurs. Les données évoluent loin des personnes auxquelles elles se réfèrent, dans des formes chiffrées étrangères au vocabulaire courant des internautes. Au cours de la donnéification, les réalités plurielles et qualitatives constitutives des identités individuelles sont réduites à des fragments calculables (Quessada, 2010), se rapportant davantage à des « dividiels » numériques, entités standardisées et interchangeables, qu'aux personnes de chair et d'os navigant sur internet (Deleuze, 1990). Les données existent donc en quelque sorte dans un univers parallèle au monde social dans lequel vivent les individus ; si ces deux espaces se correspondent et s'influencent en de nombreux points, ils sont caractérisés par

² Source : <https://netograph.io/>, consulté le 12/09/2018

des grammaires et des modes de fonctionnement propres, faisant des internautes des étrangers face aux données les concernant (Jaton & Vinck, 2016). Ceux-ci sont en effet rarement amenés à visualiser, et ne déterminent jamais, les catégories, seuils et critères à partir desquels les données prennent forme. Lorsque qu'ils sont confrontés à « leurs données », celles-ci sont souvent traitées en amont (traduites) pour leur être présentées dans une forme compréhensible, comme c'est le cas lorsqu'ils exercent leur droit d'accès garanti par le Règlement Général sur la Protection des Données (RGPD). Par ailleurs, les données exploitables sont rarement construites uniquement à partir de l'interface : différentes catégories d'informations sont le plus souvent combinées pour aboutir, dans un second temps, à des bases de données utilisables – c'est ce qu'a montré une équipe de recherche dans une vaste cartographie des données issues de Facebook (Joler & Petrovski, 2016). A titre d'exemple, sur OkCupid, la localisation des utilisateurs est inférée des quatre chiffres constituant leur adresse IP, qui ne désignent en premier lieu que la position de leur dispositif de connexion au sein de leur réseau résidentiel, attribué par leur fournisseur d'accès. Ce n'est qu'en croisant cette information avec la base du registre internet correspondant (RIPE en Europe) que les chiffres de l'adresse IP peuvent indiquer une position géographique approximative, formulée en termes de pays, de ville, et, le cas échéant, de zone urbaine. Les internautes ne connaissent généralement pas leur adresse IP, ni la signification précise des numéros qui la composent. Dans cette forme brute, les données ne correspondent pour eux à aucune réalité concrète, et ne leur sont d'aucune utilité.

La plateforme OkCupid joue ainsi un rôle prépondérant dans la construction des données qui prend place en son sein. Orientée par de forts intérêts économiques, elle met en œuvre des stratégies graphiques et techniques influençant les comportements et appréciations de ses membres, et encourageant leur exposition à la donnification. C'est à son niveau que sont déterminés non seulement les caractéristiques de l'espace de donnification, mais aussi le champ d'application de celle-ci, les moyens de sa mise en œuvre, et la forme des objets ainsi construits.

C'est au sein de ce cadre que naviguent les membres d'OkCupid. La partie suivante sera consacrée aux modalités de leur participation à la donnification dans cet espace, à leurs pratiques et à leurs perceptions. Nous comprendrons que, bien qu'acteurs essentiels d'un processus dont ils sont la pièce maîtresse, leur engagement conscient dans la donnification est structurellement limité : engagés dans des activités de nature diverses (ici, la rencontre) les individus n'appréhendent qu'une portion partielle et fluctuante du champ couvert en parallèle par cette opération. De plus, en lien avec la douceur des mécanismes de donnification faisant davantage reposer leurs possibilités d'engagement sur des

connaissances théoriques acquises par ailleurs que sur des ressentis associés aux usages, les internautes présentent des niveaux inégaux de sensibilisation et donc d'exposition à la création de données.

3. Les internautes face à la donnéification

3.1. Méconnaissance des conditions de donnéification

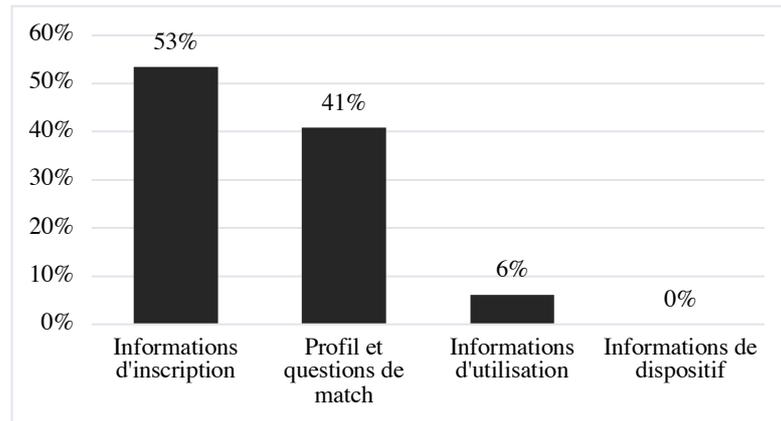
La structure économique, graphique et technique au sein de laquelle les données sont créées est à l'origine, sur OkCupid, d'une asymétrie d'engagement dans cette activité entre plateforme et utilisateurs.

Du côté de la plateforme, le pouvoir de détermination des modalités de donnéification et le caractère automatique de ces processus permettent une approche globale des bases de données ainsi constituées, stockées dans des *data centers* loués par OkCupid. Cette appréhension d'ensemble n'est en revanche pas permise aux membres du service, occupés par leurs activités de rencontre au moment de la donnéification, et ne pouvant ensuite accéder qu'avec difficulté au récapitulatif des données créées à leur sujet (l'exercice du droit d'accès est fastidieux : après une demande par courrier postal, plusieurs semaines d'attente sont nécessaires pour l'obtention du document). De ce fait, un décalage se forme entre la réalité de la création de données mise en œuvre par la plateforme et sa perception par les internautes.

La politique de confidentialité d'OkCupid distingue trois grandes catégories de données qu'elle « collecte » (contribue à créer, donc) : les informations liées au profil et aux questions de match (nom, mail, ville, âge, photos, genre, orientation sexuelle, réseaux sociaux, mode de vie, religion, ethnicité, physionomie), les informations d'utilisation (dates et lieux de connexion, messages échangés, pages visitées, historique de recherches, informations financières) et les informations concernant le dispositif utilisé par les membres (identifiant du téléphone ou de l'ordinateur, adresse IP, géolocalisation, fournisseur d'accès, qualité de connexion).

La perception des utilisateurs de ce champ recouvert par la donnéification est limitée, floue et fluctuante. Si, comme représenté sur le graphique suivant, les éléments fournis à l'inscription (mail, ville, âge, orientation sexuelle...) sont fréquemment considérés comme des données par les utilisateurs (53% d'entre eux), les traces, correspondant aux informations d'utilisation et de dispositif, le sont beaucoup moins. Seuls 6% ont conscience de partager des informations avec la plateforme au cours d'utilisations courantes (en envoyant des messages,

par exemple), et aucun utilisateur ne fait mention d'information liées à son dispositif de connexion



*Figure 1 - Fréquence de citation des différents types d'éléments donnifiés
Champ : 250 répondants à l'enquête
Source : auteure*

Deux facteurs liés permettent de comprendre cette hétérogénéité de perception : d'une part, la donnéification concerne des facettes diverses de l'identité des internautes et, d'autre part, elle est mise en œuvre à différents stades d'utilisation du service, présentant chacun des possibilités d'engagement spécifiques et inégales.

Les caractéristiques identitaires concernées par la donnéification sont loin d'être toutes équivalentes aux yeux des individus. Dans une étude consacrée à l'exposition des identités en ligne, Dominique Cardon les a classées sur un axe opposant l'être au faire, borné par les pôles de l'« identité civile » d'une part et de l'« identité agissante » d'autre part (Cardon, 2008). La première extrémité réunit des informations attachées de façon univoque et (relativement) définitive à l'individu (nom, âge, genre...) : elles sont en ce sens constitutives de ce qu'on appellera une « identité statique ». A mesure que l'on s'en éloigne, les caractéristiques identitaires deviennent de plus en plus fluctuantes, recouvrant les engagements, activités et préférences des individus.

Ces différentes composantes identitaires sont converties en données à des moments variés, correspondant à des possibilités inégales d'implication dans ces processus pour les individus. L'inscription est le temps privilégié de la donnéification de l'identité statique, et c'est aussi la seule phase à laquelle les

membres participent activement en remplissant des champs prévus à cet effet. Par la suite, les usages se focalisent de façon croissante sur l'activité primaire de rencontre à mesure que l'attention des internautes s'éloigne d'une création de données concernant davantage des éléments liés à leur identité agissante. À l'extrémité de ce continuum, les traces renseignent précisément sur les activités des individus (localisation dans l'espace, par exemple), tout en étant donnifiées d'une façon douce n'interférant pas avec leurs activités. Ces données n'offrent donc pas de « prise » visible qui leur permettraient d'être directement appréhendées comme telles (Bessy & Chateauraynaud, 1993), et elles entrent de ce fait plus difficilement dans le champ de perception de la donnification des utilisateurs.

C'est à cette lumière que l'on peut analyser le fait que de nombreux utilisateurs estiment à tort avoir limité leur exposition à la donnification. C'est le cas de Serena (23 ans, artiste), qui pense avoir transmis peu d'informations à OkCupid alors qu'elle a notamment répondu à plusieurs centaines de questions de match pouvant être considérées comme très intimes :

- Question : T'as déjà pensé aux informations que l'application a sur toi, tes données personnelles ?
- Serena : Genre mon âge, les trucs comme ça ? Non, parce que j'essaye de pas trop mettre d'informations personnelles sur OkCupid (...). Genre sur la ville j'ai mis Paris en général (...) Et j'ai mis mon âge, mais voilà parce que forcément fallait le mettre. Et d'autres informations j'essaye de pas trop mettre. Tu vois où j'ai grandi, ou les informations basiques, j'essaye de pas trop les donner.

De fait, les utilisateurs sous-estiment fréquemment la quantité d'informations exposées lors de la phase de réponse aux questions de match : comme l'illustre le graphique suivant, si les membres interrogés ont en moyenne répondu à 371 questions, leur estimation de ce chiffre ne monte qu'à 199, l'écart pouvant atteindre des proportions très importantes³.

³ Pour chaque membre interrogé, l'estimation du nombre de questions de match renseignée dans le questionnaire a été croisée avec le nombre réel visible sur chaque profil.

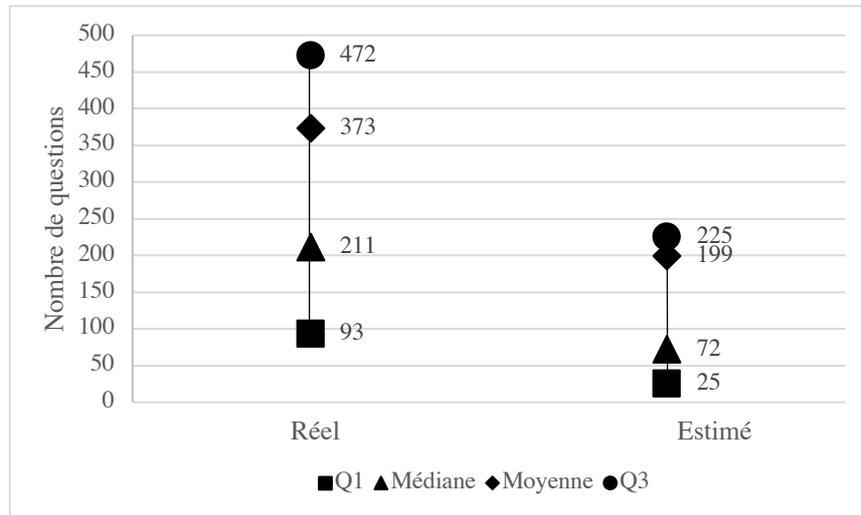


Figure 2 - Questions de match : réalité vs. perception
 Champ : 250 répondants à l'enquête
 Source : auteure

Ainsi, comme le résume le schéma suivant, si les individus ont souvent conscience de l'existence de données correspondant à leur identité statique, pour lesquelles ils ont transmis des informations à la plateforme lors de leur inscription, cette perception de la donnéification devient de plus en plus floue à mesure qu'elle s'applique à des éléments constitutifs de leur identité agissante, et qu'elle dépend de mécanismes de donnéification doux se superposant aux activités de rencontre. Ce que l'on peut appeler le « spectre de données conscientisées » des individus, c'est-à-dire l'ensemble des éléments que ceux-ci perçoivent comme ayant été convertis en données au cours de leur usage du service, est donc limité, et concentré principalement sur les éléments fournis à l'inscription – bien que, comme nous le montrerons ensuite, ce spectre n'a pas la même amplitude pour tous et toutes.

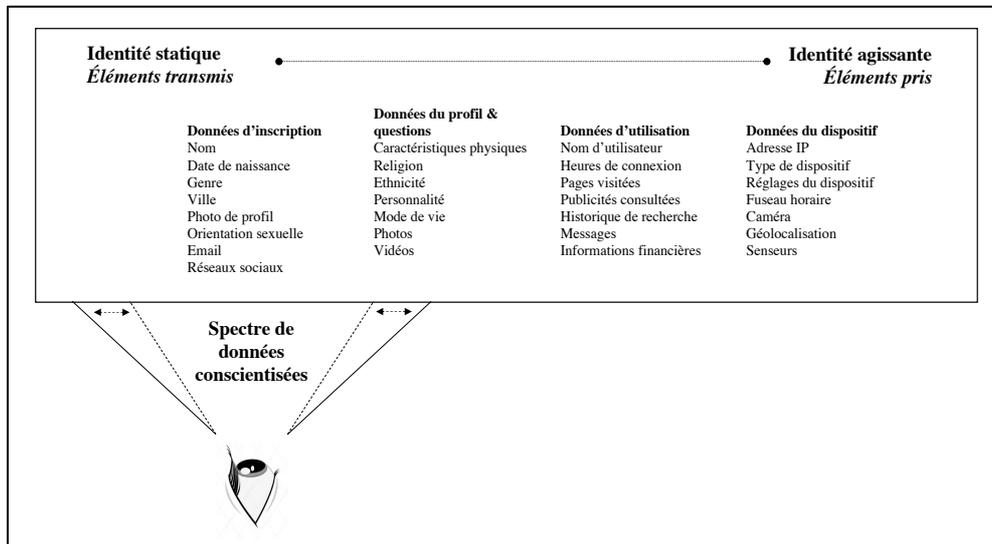


Figure 3 - Spectre des données conscientisées

3.2. Impact de la connaissance sur les comportements

A l'exception de l'inscription, la donnification se déroule donc généralement en arrière-plan de l'activité primaire de rencontre poursuivie par les utilisateurs. La douceur et l'invisibilité des mécanismes de création de données font reposer les possibilités d'engagement conscient des individus dans ces processus davantage sur des connaissances théoriques lues dans la politique de confidentialité ou acquises par ailleurs, que sur des réalités sensibles directement perceptibles à l'usage. En tant qu'utilisateur, être engagé dans la donnification, *a minima* par le biais d'un accompagnement conscient de ces processus, demande de réels efforts, dans un premier temps d'information sur les modalités de création des données, puis d'orientation de la pensée vers cet objet tout au long des usages du service.

3.2.1. Connaissance et appartenance sociale

La nécessité de cet effort a pour conséquence une différenciation des engagements des utilisateurs. Tous n'ont pas les mêmes compétences ni les mêmes motivations pour s'investir dans la donnification : des formations et

environnements différenciés conduisent à des apprentissages contrastés, et, en conséquence, à des niveaux inégaux d'exposition à la donnification.

Les niveaux de connaissance de la politique de confidentialité d'OkCupid ont été estimés au sein de notre échantillon grâce à un indicateur construit sur la base de leurs réponses plusieurs questions concernant les modalités d'usage des données par la plateforme. Les utilisateurs ont été répartis en quatre catégories : connaissance très faible, faible, intermédiaire, et expertise.

En accord avec les résultats de travaux antérieurs (Hoofnagle et al., 2010; Litman-Navarro, 2019), le niveau général de connaissance de la politique de confidentialité – très rarement lue –, et donc des modalités de création et d'utilisation des données personnelles, est bas : seuls 3% des utilisateurs interrogés peuvent être qualifiés d'experts (réponse correcte à au moins 75% des questions) tandis que la moitié dispose d'un niveau de connaissance très faible (réponse correcte 25% ou moins).

Cette ignorance est souvent consciente : 85% des personnes interrogées affirment spontanément ne pas connaître la réponse à au moins une question. Elle est liée à l'invisibilité des modalités de création et d'utilisation des données personnelles, renforcée chez certains internautes par un évitement des frictions représentées par la lecture des politiques de confidentialité les éloignant de leur activité principale sur la plateforme, comme l'explique Louise (30 ans, assistante de production) :

C'est pas très bien, hein, mais si par exemple y a une mise à jour, enfin tu sais, « on a changé de politique machin », je fais ok quoi, je vais même pas m'intéresser au truc (...). Je pense qu'ils pourraient vraiment me faire signer n'importe quoi. En fait moi je veux que ça aille vite. Et je veux que ça reste quand même un truc divertissant et en fait ça me saoule les conditions machin. Une fois j'ai essayé de les lire hein, je te promets, et j'ai commencé (...). Je me suis dit, ça serait quand même bien de lire un peu, de s'intéresser. Et j'ai dû lire... En fait j'arrivais pas à me concentrer, je pensais à autre chose tu vois, et j'ai dit oh ben tant pis. Allez, ok.

Le niveau de connaissance n'est pas homogène parmi les membres interrogés. Les personnes proches du secteur numérique présentent en particulier une sensibilisation importante aux opérations de donnification. Il y a ainsi deux fois plus de personnes travaillant dans le milieu informatique (intégration web, développement et ingénierie informatique, UX/UI design...) parmi les experts et connaisseurs intermédiaires que ce qui serait attendu en cas d'une répartition indépendante du secteur professionnel. De fait, c'est en ayant accès aux mécanismes normalement invisibles se déroulant en *back-office*, et étant intégré à des groupes portant une attention particulière à la donnification, qu'on

apprend à être vigilant à ce sujet : il faut être socialisé à ce processus pour pouvoir s'y impliquer. C'est ce que Jean, que nous avons rencontré plus haut, explique :

J'y pense assez souvent oui [à la donnéification]. Comme je te le disais tout à l'heure, je fais de la communication dans le domaine informatique (...) donc c'est des problématiques auxquelles je suis un peu sensibilisé, je lis énormément de choses, je fais de la veille dans le milieu.

La composition sociale spécifique de ces milieux fortement sensibilisés éclaire d'autres lignes de fracture dans les niveaux de connaissance. En particulier, la faible féminisation des professions informatiques, ainsi que la socialisation genrée au numérique (Collet, 2019; Dagnaud, 2018), se traduit par une méconnaissance plus importante chez les femmes des modalités de création des données personnelles. Ainsi, au sein de l'échantillon étudié, elles sont 7 fois moins nombreuses que les hommes à pouvoir être qualifiées d'expertes et, à l'inverse, leur appartenance à la catégorie « connaissance très faible » est deux fois plus importante que ce qui aurait été attendu en cas d'une répartition indépendante du genre.

3.2.2. Connaissance inégales : expositions inégales

L'inégalité de connaissance et de compréhension des modalités de donnéification se traduit par des expositions différentes des individus à ce processus. Notre enquête nous conduit à formuler l'hypothèse d'une relation non-linéaire unissant ces deux variables, suivant quatre modalités distinctes illustrées sur la représentation schématique suivante :

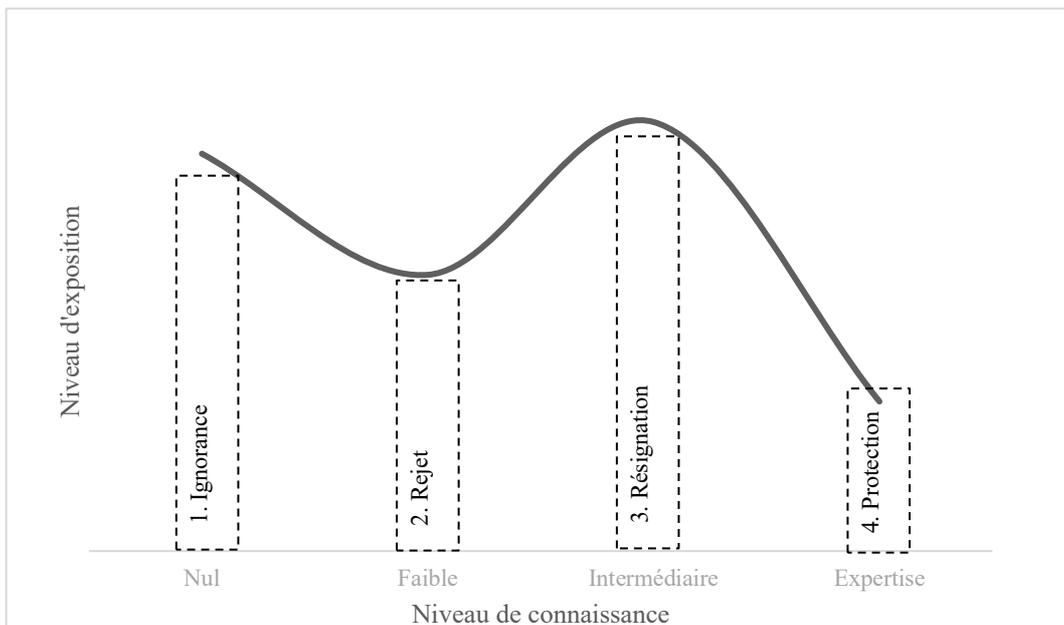


Figure 4 – Représentation schématique de l'exposition à la donnéification en fonction du niveau de connaissance

Ces différentes configurations de l'articulation connaissance-exposition peuvent se définir comme suit :

- 1- **Ignorance** : initialement et en l'absence de toute connaissance, l'exposition des individus se situe à un niveau « neutre », non influencé par la prise en compte de cet élément.
- 2- **Rejet** : l'acquisition des premières briques de connaissance est liée à une diminution de la quantité d'informations rendues disponibles sur l'interface. On peut assimiler ce phénomène à une phase de rejet de la surveillance numérique, comme l'explique Sony (26 ans, développeur web), un utilisateur se situant dans cette catégorie :

Sony : Moi, sur internet, j'évite de tout balancer. Quand on me demande des trucs sexuels sur un site de rencontres, je réponds pas.

Question : Ça a toujours été comme ça ou il y a des choses qui t'ont fait aller dans ce sens ?

Sony : C'est une évolution, enfin c'est pas vraiment un déclic, c'est vraiment une évolution dans mon état de pensée (...). C'est à force de voir toutes les histoires de personnes qui se sont fait voler des données, de voir même techniquement comment tu peux récupérer des données, comment c'est utilisé par les grandes sociétés, (...) je préfère rester prudent avec tout ça.

- 3- **Résignation** : ce n'est que dans un troisième temps que l'exposition augmente parallèlement à la connaissance. On peut supposer qu'à partir d'un certain seuil de conscientisation, les individus se sentent « résignés à l'inévitabilité de la surveillance et au pouvoir des entreprises collectant leurs données » (Turow et al., 2015). La protection de surface mise en œuvre dans la phase de rejet semble alors vaine, et, puisque qu'il est difficile de se couper de tout service numérique, l'exposition ne dépend plus de la préoccupation suscitée par la donnéification : elle redevient similaire à celle des individus « ignorants ». Antoine (25 ans, vendeur Apple), un utilisateur qui peut être qualifié de résigné, parle de cette exposition comme d'un « jeu » dans lequel il serait contraint de s'impliquer :

Question : Tu as déjà pensé aux infos dont l'app dispose à ton sujet ?

Antoine : Je connais. J’y ai pensé. C’est un jeu, dans le sens où de toutes façons tout le monde a tes données sur n’importe quelle interface. T’as Facebook on a tes données. T’as LinkedIn on a tes données professionnelles. T’as OkCupid on a tes données plus perso. (...) Je fais des achats sur Amazon ils ont mes données d’achat (...). C’est soit tu joues le jeu et tu te poses pas trop ce genre de questions (...), soit tu te poses les questions et du coup t’arrêtes de faire tout, quoi. Si tu veux pas que tes données soient publiées en ligne ben tu peux pas faire grand-chose malheureusement aujourd’hui. C’est un jeu qu’il faut jouer, c’est un jeu qu’il faut accepter de jouer.

4- **Protection** : enfin, les rares experts interrogés présentent un niveau d’exposition bien moindre à celui de l’ensemble de l’échantillon. Un tel niveau de connaissance semble en effet s’accompagner d’une conscientisation aiguë des possibilités d’exploitation des données par la plateforme, et de l’acquisition de compétences techniques nécessaires à une non-exposition efficace.

C’est le cas de Seguromartiger (26 ans, étudiant), doctorant dans le domaine de la cyber sécurité, qui a très peu rempli son profil sur OkCupid (nombre minimum de questions de match répondues, pas de photo, indication d’un pseudo plutôt que du nom) :

Je m’y connais bien en informatique et en sécurité, je sais que c’est souvent un mauvais plan de donner trop d’informations (...). Même si tu réponds juste par oui ou non [à des questions de match], tu les renseignes déjà beaucoup. Tu sais pas comment les apps vont utiliser tes données. Et même si c’est des gentils ils peuvent aussi avoir des gros problèmes de sécurité (...) Donc oui, je protège ma localisation et mon profil est vide pour ne pas diffuser trop de données personnelles.

Tirée d’une enquête par entretiens et questionnaire, l’hypothèse de la relation non-linéaire unissant connaissance et exposition aux processus de donnéification pourrait être testée à plus grande échelle par le biais d’une enquête quantitative orientée vers l’association de ces deux variables. Ce schéma analytique permet en tous cas de comprendre que participer à la donnéification, c’est avant tout s’en protéger. Lorsque la donnéification est prise en compte par les membres d’OkCupid dans la définition de leurs actions, c’est rarement pour coopérer positivement avec la plateforme dans la réalisation de cette tâche, mais plutôt pour limiter l’ampleur du champ couvert par ce processus.

Ainsi, les caractéristiques de la donnéification sur OkCupid, largement déterminées par la plateforme, impactent les perceptions et les rapports que les individus peuvent entretenir avec ce processus. La douceur des mécanismes de collecte ainsi que la forme de l'interface limitent la conscience que les internautes peuvent avoir de la création de données, et conditionne leurs possibilités d'engagement à l'acquisition extérieure de connaissances. Les données font donc l'objet d'appréhensions et de définitions variables. Leur existence est intrinsèquement imbriquée aux activités menées par les individus sur les plateformes, par lesquelles elles émergent et qu'elles conditionnent en retour.

4. Conclusion

L'objectif de cet article était de proposer une appréhension dynamique et contextuelle de la notion de données personnelles, en replaçant ces objets informationnels dans leur cadre de construction, celui de la rencontre entre plateforme et utilisateurs. Nous avons cherché à déconstruire une certaine appréhension des données personnelles les considérant comme des objets autonomes, possédés par les internautes qui choisiraient de les transférer aux plateformes en échange d'un service, pour les aborder plutôt comme la résultante d'un processus complexe de conversion d'éléments constitutifs des identités individuelles en informations numériques standardisées. La prise en compte du cadre dans lequel cette donnéification a lieu permet d'identifier la place occupée par les différents acteurs de ce processus, et de comprendre à cette lumière les perceptions et comportements des utilisateurs.

Dans le cas d'OkCupid, les entreprises propriétaires, pour qui l'exploitation des données personnelles générées sur leurs services constitue une source de revenus majeure, ont conçu une interface personnalisée et ludifiée encourageant l'exposition des membres à la donnéification. C'est également à ce niveau que sont configurés et mis en œuvre les moyens techniques de création des données dont la douceur permet de se passer de l'engagement conscient des utilisateurs dans cette activité. Les caractéristiques de l'espace de donnéification ainsi construit contribuent à focaliser l'attention des membres sur l'activité de rencontre, et à éloigner de leur perception les opérations de donnéification se déroulant simultanément en arrière-plan. En conséquence, les utilisateurs ont toujours une conception partielle du champ couvert par la donnéification, qui varie au fil des usages et des apprentissages. La forme spécifique d'OkCupid fait reposer la prise en compte de la donnéification davantage sur des connaissances

théoriques acquises en-dehors de la plateforme que sur des sensations liées à l'usage, connaissances face auxquelles les internautes sont loin d'être égaux. Cette appréhension différenciée est à l'origine de possibilités de participation ou de résistance inégales à la donnification, et donc de niveaux différenciés d'exposition en fonction des milieux d'appartenances et des sensibilisations. Il ressort ainsi de cette enquête que les données ne constituent pas un ensemble homogène, mais représentent au contraire des réalités plurielles liées la nature des objets auxquels elles se rattachent, à leur processus de création et aux appréhensions que les individus en ont.

Une telle configuration n'est pas l'apanage d'OkCupid. Les interactions prenant place dans les espaces numériques, intermédiées par un grand nombre de dispositifs techniques et d'acteurs tiers, reposent en grande partie sur la circulation de données personnelles. Objets que l'on pourrait qualifier de « frontières » (Star & Griesemer, 1989), les données permettent d'articuler l'ensemble hétérogène d'éléments matériels et immatériels sous-tendant l'existence des communautés et des activités numériques. La définition de leur forme et de leurs contours, dans laquelle la position surplombante des plateformes joue un rôle déterminant, reste conditionnée à leur contexte de création, et à la participation – active ou non – des différents internautes engagés dans les activités offertes sur leurs interfaces.

Bibliographie

- Acquisti, A. (2011). Les comportements de vie privée face au commerce électronique : Une économie de la gratification immédiate. *Réseaux*, vol. 167, n°3, p. 105-130.
- Akrich, M. (1990). De la sociologie des techniques à une sociologie des usages. *Techniques et Culture*, n°16, p. 83-110.
- Barnes, S. B. (2006). A privacy paradox: Social networking in the United States. *First Monday*, vol. 11, n° 9.
- Bastard, I., Cardon, D., Charbey, R., Cointet, J.-P., & Prieur, C. (2017). Facebook, pour quoi faire ? *Sociologie*, vol 8, n°1, p. 57-82.
- Bessy, C., & Chateauraynaud, F. (1993). Les ressorts de l'expertise. Epreuves d'authenticité et engagement des corps. *Les objets dans l'action. De la maison au laboratoire* Paris, Editions EHESS, p. 141-164.
- Beuscart, J.-S., Coavoux, S., & Maillard, S. (2019). Les algorithmes de recommandation musicale et l'autonomie de l'auditeur. *Réseaux*, vol 1, n°213, p. 17-47.
- Bujlow, T., Carela-Espanol, V., Lee, B. R., & Barlet-Ros, P. (2017). A Survey on Web Tracking : Mechanisms, Implications, and Defenses. *Proceedings of the IEEE*, vol. 105, n°8, p 1476-1510.
- Callon, M. (2013). Qu'est-ce qu'un agencement marchand ? *Sociologie des agencements marchands*. Paris, Presses des Mines, p. 325-440.
- Cardon, D. (2008). Le design de la visibilité. *Réseaux*, vol. 152, n°6, p. 93-137.
- Cardon, D. (2019). *Cultures Numériques*. Paris, Presses de Sciences Po.
- Chaves Ferreira, B., Jourdain, A., & Naulin, S. (2018). Les plateformes numériques révolutionnent-elles le travail ? *Réseaux*, vol 6, n°212, p. 85-119.
- Collet, I. (2019). *Les oubliées du numérique*. Paris, Le Passeur.
- Crain, M. (2018). The limits of transparency : Data brokers and commodification. *New Media and Society*, vol 20, n°1, p. 88-104.
- Dagnaud, M. (2018). Internet, une passion masculine. *Le Débat*, vol 3, n°200, p. 123-142.
- Deleuze, G. (1990). Post-scriptum sur les sociétés de contrôle. *Pourparlers*. Paris, Editions de Minuit, p. 240-247.
- Denis, J. (2018). *Le travail invisible des données. Elements pour une sociologie des infrastructures scripturales*. Paris, Presses des Mines.
- Deterding, S. (2012). Gamification : Designing for motivation. *Interactions*,

- vol 19, n°4, p. 14-17.
- Hallinan, B., & Striphas, T. (2016). Recommended for you : The Netflix Prize and the production of algorithmic culture. *New Media and Society*, vol. 18, n°1, p. 117-137.
- Hoofnagle, C. J., King, J., Li, S., & Turow, J. (2010). How Different are Young Adults from Older Adults When it Comes to Information Privacy Attitudes and Policies? *SSRN Electronic Journal*.
- Hughes, T. P. (1983). *Networks of power. Electrification in Western Society, 1880-1930*. Baltimore & London, The Johns Hopkins University Press.
- Jaton, F., & Vinck, D. (2016). Processus frictionnels de mises en bases de données. *Revue d'anthropologie des connaissances*, vol. 10, n° 4, p. 489-504.
- Joler, V., & Petrovski, A. (2016). *Facebook Algorithmic Factory. Immaterial Labour and Data Harvesting*, <https://labs.rs/en/facebook-algorithmic-factory-immaterial-labour-and-data-harvesting/>
- Jouët, J. (2000). Retour critique sur la sociologie des usages. *Réseaux*, vol. 18, n° 100, p. 487-521.
- Just, N., & Latzer, M. (2017). Governance by algorithms : Reality construction by algorithmic selection on the Internet. *Media, Culture and Society*, vol. 39 n°2.
- Licoppe, C. (2007). From Interpersonal Communication to Epistemic Communities : ICT Development and the Paradigm of Distribution. *Hermès, La Revue*, vol. 47, n°1, 59-67.
- Litman-Navarro, K. (2019). We Read 150 Privacy Policies. They Were an Incomprehensible Disaster. <https://www.nytimes.com/interactive/2019/06/12/opinion/facebook-google-privacy-policies.html>
- Loveluck, B. (2015). *Réseaux, libertés et contrôle : Une généalogie politique d'internet*. Paris, Armand Colin.
- Lyon, D., Haggerty, K. D., & Ball, Kirstie. (2012). *Routledge handbook of surveillance studies*. New-York, Routledge.
- MacKenzie, D. (1990). *Inventing Accuracy. A historical sociology of nuclear missile guidance*. Cambridge, The MIT Press.
- Marx, G. T. (2006). Soft Surveillance. Hey Buddy can you spare a DNA? New surveillance technologies and the growth of mandatory volunteerism in collecting personal information. *Lex electronica*, vol 10 n°3.
- Möller, J., Trilling, D., Helberger, N., & van Es, B. (2018). Do not blame

- it on the algorithm : An empirical assessment of multiple recommender systems and their impact on content diversity. *Information Communication and Society*, vol. 21, n°7, p. 959-977.
- Nissenbaum, H. (2011). A contextual approach to privacy online. *Daedalus*, vol. 140, n°4, p. 32-48.
- Quessada, D. (2010). De la sousveillance. *Multitudes*, vol. 40, n°1, p. 54-59.
- Reigeluth, T. (2018). La prédiction algorithmique comme activité sociale. *Réseaux*, vol. 211, n°5, 35-67.
- Rochelandet, F. (2010). *Économie des données personnelles et de la vie privée*. La Paris, Découverte.
- Rosenblat, A. (2018). *Uberland. How algorithms are rewriting the rules of work*. Berkeley, University of California Press.
- Rouvroy, A., & Berns, T. (2010). Le nouveau pouvoir statistique. *Multitudes*, vol. 40, n°1, p. 88-103.
- Rudder, C. (2014). *Datachism : Love, Sex, Race and Identity. What our Online Lives tell us about our offline Selves*. New-York, Broadway Books.
- Sadin, É. (2010). Le nouveau paradigme de la surveillance. Cerner l'humain par l'entrelacs du marketing et de la sécurité. *Multitudes*, vol. 40, n°1, p. 54-59.
- Star, S. L., & Griesemer, J. R. (1989). Institutional Ecology, « Translations » and Boundary Objects : Amateurs and Professionals in Berkeley's Museum of Vertebrate Zoology. *Social Studies of Science*, vol. 19, n°3, p. 387-420.
- Star, S. L., & Strauss, A. (1999). Layers of Silence, Arenas of Voice : The Ecology of Visible and Invisible Work. *Computer Supported Collaborative Work*, vol. 8, n°1-2, p. 9-30.
- Tubaro, P., & Casilli, A. (2017). Enjeux sociaux des Big Data. *Les Big data à découvert*. Paris, CNRS Editions, p. 292-293.
- Turow, J., Hennessy, M., & Draper, L. (2015). The Tradeoff Fallacy : How Marketers are Misrepresenting American Consumers and Opening Them Up to Exploitation. *Annenberg School of Communication*, n°24.
- Van Dijck, J. (2014). Datafication, dataism and dataveillance : Big data between scientific paradigm and ideology. *Surveillance and Society*, vol. 12, n°2, p. 197-208.
- Zuboff, S. (2015). Big other : Surveillance capitalism and the prospects of an information civilization. *Journal of Information Technology*, vol. 30, n°1, p. 75-89.

