



Méthodes de recherche en contexte numérique

Une orientation qualitative

Sous la direction de MÉLANIE MILLETTE, FLORENCE MILLERAND,
DAVID MYLES et GUILLAUME LATZKO-TOTH

MÉTHODES DE RECHERCHE EN CONTEXTE NUMÉRIQUE

Sous la direction de
Mélanie Millette, Florence Millerand,
David Myles et Guillaume Latzko-Toth

MÉTHODES DE RECHERCHE EN CONTEXTE NUMÉRIQUE

Une orientation qualitative

Les Presses de l'Université de Montréal

Catalogage avant publication de Bibliothèque et Archives nationales du Québec et Bibliothèque et Archives Canada

Titre: Méthodes de recherche en contexte numérique : une orientation qualitative / [sous la direction de] Mélanie Millette, Florence Millerand, David Myles et Guillaume Latzko-Toth. Noms: Millette, Mélanie, 1980- éditeur intellectuel. | Millerand, Florence, éditeur intellectuel. | Myles, David, 1987- éditeur intellectuel. | Latzko-Toth, Guillaume, 1972- éditeur intellectuel. | Congrès de l'ACFAS (84e : 2016 : Université du Québec à Montréal) Collections: PUM.

Description: Mention de collection: PUM | Textes présentés lors d'un colloque tenu en 2016 dans le cadre du 84e Congrès de l'ACFAS. | Comprend des références bibliographiques et un index.

Identifiants: Canadiana (livre imprimé) 20200076868 | Canadiana (livre numérique) 20200076876 | ISBN 9782760642492 (couverture souple) | ISBN 9782760642508 (PDF) | ISBN 9782760642515 (EPUB)

Vedettes-matière: RVM: Médias sociaux—Recherche—Méthodologie. | RVM: Sciences sociales—Recherche—Informatique. | RVM: Recherche qualitative—Méthodologie.

Classification: LCC HM742.M48 2020 | CDD 302.23/10723—dc23

Mise en pages: Folio infographie

Dépôt légal: 3^e trimestre 2020

Bibliothèque et Archives nationales du Québec

© Les Presses de l'Université de Montréal, 2020

Cet ouvrage a été publié grâce à une subvention de la Fédération des sciences humaines de concert avec le Prix d'auteurs pour l'édition savante, dont les fonds proviennent du Conseil de recherches en sciences humaines du Canada.

Les Presses de l'Université de Montréal remercient de son soutien financier la Société de développement des entreprises culturelles du Québec (SODEC).

Financé par le gouvernement du Canada
Funded by the Government of Canada

Canada

SODEC
Québec

IMPRIMÉ AU CANADA

Préface

Dominique Pasquier

Qu'on ne s'y trompe pas : ce livre n'est pas un manuel de méthodologie. Il repose même sur la démarche inverse : au lieu de dicter des consignes, il parle de doutes, d'essais, de limites. C'est ce qui lui donne toute sa valeur tant il est rare que les universitaires acceptent d'ouvrir la boîte noire de leurs enquêtes pour évoquer, dans la plus grande transparence, tous les problèmes qui jalonnent un travail de terrain. On comprend, en le lisant, que le numérique, comme outil, comme objet ou comme terrain – car ces trois dimensions sont abordées –, ne facilite pas la conduite de recherches en sciences sociales. Mais, en même temps, que de mondes sociaux nouveaux à explorer, que de « bricolages » méthodologiques originaux à concevoir – pour parler comme De Certeau –, que de ressources inédites à utiliser ! C'est le grand intérêt du présent ouvrage, et, du même coup, sa réelle qualité pédagogique : chaque chapitre illustre à sa manière le bouleversement que la numérisation de la vie sociale impose aujourd'hui aux sciences sociales.

Ce bouleversement a souvent été analysé du côté des *big data*, ces données massives qui revendiquent d'échapper aux impératifs d'échantillonnage et de représentativité, et semblent – et le verbe « sembler » prend ici tout son sens – périmer, du fait de leur volume et de leur immédiateté, les enquêtes traditionnelles par questionnaire (Bastin et Tubaro, 2018). Ici, c'est le renouvellement des méthodes qualitatives qui est au cœur du questionnement. D'une part, ces méthodes continuent d'être un complément indispensable aux recherches fondées sur des données massives (en permettant de les « densifier », comme on l'explique au chapitre 10), mais surtout, la mutation qu'elles ont subie avec l'arrivée de nouvelles méthodes et de nouveaux terrains de recherche est particulièrement foisonnante et intéressante à suivre.

Les questions ne manquent pas. Peut-on encore parler d'ethnographie quand les observations se font en ligne, et souvent à l'insu des personnes mêmes ? Peut-on faire de la sociologie quand les individus interrogés se cachent derrière l'anonymat ou le pseudonymat ? Quel degré de généralité accorder à des interactions qui sont dominées par un très petit nombre de « locuteurs dominants » qui occupent pratiquement tout le terrain (Akrich, 2012) ? Comment analyser des échanges qui perturbent les principes de paires adjacentes et d'ouverture/clôture mis en évidence par l'analyse d'une conversation ? Quel sens donner à la notion de « communauté » quand les engagements dans l'action sont aussi hétérogènes ? Comment gérer la forte réversibilité des positions de producteur et de récepteur qui caractérise les mondes en ligne ? Peut-on fonder son analyse sur un seul site alors que les personnes interrogées naviguent sur une multiplicité de sites en ligne et que c'est cette circulation de l'un à l'autre qui a du sens pour elles ? Comment découper un objet de recherche cohérent au sein d'un univers sans frontières fixes ?

On pourrait allonger la liste à l'infini. Le numérique emprunte à la fois aux règles de la communication interpersonnelle et à celles des médias pour créer un système d'échanges d'un type nouveau, souvent fondés sur des relations à plusieurs à distance, synchrones comme asynchrones, qui entremêlent étroitement conversation et publication et laissent des traces (Beaudouin, 2002). Les règles habituelles de l'échange y sont perturbées dans une lutte permanente pour la captation de l'attention, comme en témoigne par exemple la part très importante de messages qui restent sans réponse sur les forums (Beaudouin et Velkovska, 1997). En outre, comme le souligne la direction de l'ouvrage, la circulation des images y tient une place sans cesse plus grande. Autant dire que la transformation n'est pas anodine et qu'elle pose un véritable défi aux sciences sociales.

Pourtant, il ne faut pas sombrer dans l'illusion du « tout nouveau ». Le monde en ligne est intrinsèquement dépendant de ce qui se passe dans le monde hors-ligne, au point que plusieurs scientifiques n'envisagent plus d'étudier l'un sans l'autre. De nombreux chapitres du présent ouvrage défendent cette option, sous des formes différentes, soit en suivant un groupe d'individus à la fois dans la vie réelle et dans leurs univers de navigation, soit en confrontant les personnes interrogées à leurs traces numériques (« entretien de confrontation »), soit encore en leur demandant d'expliquer leur parcours en ligne (« visite commentée »). Les résultats sont

très concluants et permettent de lever bien des ambiguïtés quant aux motivations réelles des pratiques en ligne. Le numérique n'abolit pas non plus les barrières sociales contrairement aux discours utopistes que pouvaient tenir les pionniers et pionnières d'Internet. Ce sont les individus les plus diplômés qui participent le plus, en produisant des contenus créatifs, en intervenant sur les sites de discussion, en réagissant aux articles de presse (Schradié, 2011). Ceux qui ont un faible capital scolaire instaurent des usages modestes et pragmatiques : Internet ne leur ouvre pas le monde, il leur permet de resserrer leurs liens familiaux ; c'est une source importante d'apprentissage et d'information, mais en contournant autant que possible l'écrit au profit des tutoriels vidéo ; ils font circuler beaucoup de contenus trouvés en ligne mais n'en produisent pas eux-mêmes (Pasquier, 2018). Bref, on aurait tort de croire que le monde numérique échappe aux rigidités de la vie sociale, même si l'occultation des positions sociales en ligne du fait de l'anonymat rend le travail plus difficile.

L'ethnographie en ligne et les recherches sur les traces numériques constituent une part importante de l'ouvrage. Il est vrai que ce champ de recherche est particulièrement riche en nouvelles questions et nouveaux objets et outils : articulation entre ethnographie en ligne et hors-ligne, combinaison des méthodes ethnographiques, articulation entre ethnographie en ligne et analyse de discours ou sociologie des usages, analyses des mises en scène de soi sur YouTube (étudiées dans deux chapitres sur les pratiques de visionnage en ligne adolescentes), observation des mobilités individuelles en vidéo-ethnographie avec des lunettes caméra... Les nouvelles pratiques d'enquête ont particulièrement bousculé la discipline, au point que nombre d'ethnologues refusent de considérer ces nouvelles perspectives comme faisant réellement partie de leur communauté savante. L'ethnographie en ligne, on le sait, favorise les situations d'observation invisible : c'est là une rupture avec la tradition de la discipline qui repose sur le contact étroit et prolongé entre la personne qui observe et celles qui sont observées. Cependant, comme le rappelle Christine Hine (dans le chapitre 4), « l'ethnographie en ligne en tant que champ de recherche distinct a contribué à montrer la nécessité de jeter un regard neuf sur les questions d'éthique », notamment parce qu'elle tend à positionner l'ethnologue davantage du côté de la participation que de l'observation (surtout dans les cas, extrêmes, d'autoethnographie) et que

beaucoup de recherches sont désormais menées à l'insu des personnes observées. Cette question éthique revient d'ailleurs dans plusieurs chapitres du livre : dans les textes portant sur l'analyse des « traces numériques » où l'on présente l'entretien d'autoconfrontation comme une manière de valider la participation de la personne interrogée à la recherche, mais aussi dans un chapitre tout à fait fascinant sur les métadonnées d'images dont l'analyse pose effectivement d'importants problèmes déontologiques. Tout se passe comme si la relative facilité d'accès aux données qu'offre le numérique ou l'adoption de méthodes « discrètes », voire clandestines, provoquait une certaine culpabilité dans le milieu de la recherche. A-t-on le droit de travailler sur des matériaux qui sont disponibles sans que les personnes qui les génèrent en soient averties ? Il semble qu'à cette question, il existe des réponses adaptées à chaque cas et qu'il serait dommageable pour l'avenir des recherches de devoir adopter une règle unique. Cela risque pourtant d'être le cas si l'on en juge par le durcissement actuel des chartes d'enquête universitaires.

Alors, en définitive, que retenir de ce livre ? Qu'il fait œuvre de pionnier en traitant de la question des méthodes par le biais de celles des objets de recherche, mettant ainsi en acte ce que rappelaient les auteurs du *Métier de sociologue* il y a cinquante ans : il n'y a pas de bonne ni de mauvaise méthode, il y a des méthodes plus ou moins bien appropriées à l'objet de la recherche (Bourdieu, Chamboredon et Passeron, 1968). Il faut retenir que les personnes ayant contribué à l'ouvrage ont vraiment joué le jeu, en acceptant de faire part de leurs doutes et de leurs difficultés autant que de leurs innovations et découvertes. Il faut aussi prendre en compte que l'ouvrage rend hommage aux méthodes qualitatives et à leur incomparable capacité à traiter à grain fin des phénomènes complexes. Mais surtout, il promeut une règle que tout chercheur et toute chercheuse devrait suivre : la combinaison des méthodes. Loin de s'opposer, méthodes traditionnelles et méthodes liées aux outils numériques se portent mutuellement, et sont aussi complémentaires que le sont depuis toujours les méthodes quantitatives et qualitatives. Cette leçon, ô combien utile, devrait être considérée comme une clé de lecture du présent ouvrage.

Références

- AKRICH, Madeleine, « Les listes de discussion comme communautés en ligne : outils de description et méthodes d'analyse », *Papiers de recherche du CSI*, Centre de sociologie de l'innovation, 2012.
- BASTIN, Gilles et Paola TUBARO, « Le moment *big data* des sciences sociales », *Revue française de sociologie*, vol. 59, n° 3, 2018, p. 375-394.
- BEAUDOUIN, Valérie, « De la publication à la conversation. Lecture et écriture électroniques », *Réseaux*, vol. 116, n° 6, 2002, p. 199-225.
- BEAUDOUIN, Valérie et Julia VELKOVSKA, « Constitution d'un espace de communication sur Internet (forums, pages personnelles, courrier électronique...) », *Réseaux*, n° 97, 1999, p. 121-177.
- BOURDIEU, Pierre, Jean-Claude CHAMBOREDON et Jean-Claude PASSERON, *Le métier de sociologue*, Mouton, 1968.
- PASQUIER, Dominique, *L'Internet des familles modestes. Enquête dans la France rurale*, Presses des Mines, 2018.
- SCHRADIE, Jen, « The digital production gap: The digital divide and Web 2.0 collide », *Poetics*, vol. 39, n° 2, 2011, p. 145-168.

Remerciements

L'équipe qui a dirigé cet ouvrage tient à témoigner sa profonde gratitude à son assistant éditorial, Rémi Toupin. La qualité de ce livre doit énormément au travail qu'il a accompli, travail gigantesque autant que rigoureux et diligent, tant dans la communication avec les auteurs et autrices que dans la préparation du manuscrit aux différentes étapes de son élaboration. La genèse de cet ouvrage est aussi tributaire de la tenue du colloque « Enjeux méthodologiques de la recherche sur les usages d'Internet et des technologies numériques » qui a eu lieu à l'Université du Québec à Montréal, les 11 et 12 mai 2016, sous l'égide et avec le soutien de l'Association francophone pour le savoir (Acfas). Organisé par David Myles, Mélanie Millette et Ghada Touri, ce colloque a réuni plusieurs futurs contributeurs et contributrices de cet ouvrage. Nous tenons également à remercier Christine Hine et les éditions SAGE pour la permission de traduire et de présenter un texte inédit en français. Enfin, nous sommes reconnaissants à l'équipe des PUM et particulièrement à Patrick Poirier pour son soutien enthousiaste et son accompagnement dans ce projet, ainsi qu'aux personnes ayant effectué l'évaluation externe de l'ouvrage, pour leurs commentaires inspirants et constructifs.

Introduction

*Mélanie Millette, Florence Millerand,
David Myles, Guillaume Latzko-Toth*

La méthodologie constitue un sujet passionnant pour ceux et celles qui s'intéressent à des problèmes complexes et émergents nécessitant un rapport créatif à l'enquête. C'est peut-être pourquoi la recherche menée en contexte numérique a eu tendance à insister sur les enjeux relatifs à la méthode. En effet, les chercheurs et chercheuses qui ont pris acte de l'émergence de terrains et d'objets numériques et qui ont voulu adapter les méthodes existantes ou en établir de nouvelles pour pouvoir les étudier ont rapidement dû expliquer ces méthodes en raison de leur caractère inédit, original ou innovant, voire à cause des incompréhensions et des méfiances qu'elles peuvent susciter.

Le développement des médias et technologies numériques opère d'importantes mutations – et soulève autant de résistances – dans les sciences sociales et humaines francophones, notamment en ce qui a trait à l'accès aux données numériques, à la création d'outils adaptés à leur analyse, ainsi qu'à la recomposition des frontières disciplinaires qui en découlent (Baya-Laffite et Benbouzid, 2017). Cet ouvrage collectif se situe ainsi dans le sillage de travaux pionniers, notamment en sociologie des médias et des usages des technologies (Jouët, 2000; Proulx, 2005), qui ont mis en œuvre dès les années 1980 des méthodes impliquant l'analyse fine des pratiques se déployant sur et avec des technologies de communication, ainsi que la description critique et minutieuse des dispositifs (Beuscart *et al.*, 2016). Nous élargissons également ce travail en prenant en compte la matérialité propre de certains dispositifs numériques et les enjeux éthiques abondamment traités dans les milieux anglo-saxons (Ananny

et Crawford, 2016; Luka et Millette, 2018) ainsi que dans quelques ouvrages francophones récents (Barats, 2013; Beuscart *et al.*, 2016).

Si les enjeux épistémologiques, pratiques et éthiques de la méthodologie en contexte numérique font l'objet de conversations soutenues depuis le tournant du siècle, une recension rapide des écrits scientifiques révèle qu'on a surtout produit ces contributions dans les milieux anglophones de la recherche, tout particulièrement dans le champ des *Internet studies*. C'est à la nécessité de mener une telle réflexion en langue française que s'efforce de répondre le présent ouvrage. En mettant l'accent sur la recherche qualitative, nous souhaitons aborder de manière inédite la méthodologie et les méthodes mobilisant le numérique comme objet, outil, ou terrain d'enquête.

La méthodologie : ouvrir la boîte noire

La méthodologie a une signification double. Dans son sens usuel, elle s'intéresse aux méthodes de recherche et rend compte de l'articulation entre les différentes étapes permettant d'effectuer le travail scientifique. Autrement dit, dans cette acception courante, elle s'intéresse au « comment faire » : aux décisions, opérations et analyses menées pour cerner et expliquer un phénomène. Or, au-delà des procédures permettant de réaliser la recherche, la méthodologie informe également le positionnement intellectuel des scientifiques, dans la mesure où la sélection de nos outils méthodologiques s'arrime à nos postures épistémologique et axiologique. Cette composante plus « méta » de la méthodologie concerne les manières dont sont constitués les savoirs dans l'ensemble des disciplines scientifiques. Comprises ainsi, épistémologie et méthodologie sont intrinsèquement liées : la position épistémologique éclaire la façon dont est envisagé le rapport aux personnes et phénomènes étudiés, ainsi que la sélection des méthodes utilisées pour mener la cueillette et l'analyse des données.

Ces deux dimensions de la méthodologie – pratique et épistémologique – sont abordées dans le présent ouvrage. Collectivement, nous avons voulu ouvrir la « boîte noire » de la méthodologie en contexte numérique en expliquant les enjeux épistémologiques des méthodes de collecte et d'analyse de données, ainsi qu'en soulignant leurs implications, leurs potentialités et leurs limites – le plus souvent à partir d'exemples de recherches récentes ou en cours.

L'expression « contexte numérique » renvoie à des contextes d'activité dans lesquels les médias et les technologies numériques occupent une place centrale et constituent une infrastructure essentielle pour les phénomènes, pratiques et interactions sociales sur lesquels porte la recherche. Il existe de moins en moins de contextes d'activité qui échappent à cette définition, tant « l'imbrication » d'Internet dans le tissu de la vie quotidienne se fait sentir dans toutes les sphères des sociétés contemporaines. De la même manière, les travaux en communication et en études d'Internet évitent le préfixe « cyber » depuis plusieurs années, parce qu'il reproduit un clivage factice des espaces et des expériences (le « cyber » versus la « vraie vie ») qui ne correspond pas à l'enchevêtrement observé en recherche. Parler de contexte numérique permet justement de dépasser le clivage en ligne/hors-ligne afin d'embrasser la réalité mixte, subtile et texturée des pratiques et des interactions.

Les mutations observées au cours des vingt dernières années, tant du côté des usages numériques participant des phénomènes à l'étude que des capacités à les observer et à les analyser, ouvrent de nombreuses possibilités en recherche (Millerand *et al.*, 2010). Les chapitres réunis dans cet ouvrage témoignent de celles-ci en mobilisant une série d'enquêtes et d'approches à prédominance qualitative. Dans le contexte contemporain des données massives (*big data*), notre ouvrage réaffirme la pertinence des approches qualitatives en misant notamment sur l'analyse de données denses et riches (*thick data*) pour comprendre l'émergence de phénomènes sociaux contemporains (Latzko-Toth *et al.*, 2017). Les chapitres réunis ici s'inscrivent dans une perspective compréhensive visant à rendre compte de l'expérience sociale des individus en lien avec le numérique, tout en interrogeant les implications soulevées par les conditions de la recherche empirique menée dans ce contexte.

Or, s'il s'inscrit dans une orientation qualitative, cet ouvrage ne s'oppose pas *de facto* aux méthodes quantitatives. Au contraire, nous invitons le lectorat à demeurer alerte face aux discours consistant à réifier les divisions traditionnellement établies entre les méthodes qualitatives et quantitatives.

Nous estimons que toute méthode – qu'elle soit qualitative, quantitative, mixte ou nativement numérique – peut être fertile pourvu qu'elle s'inscrive dans une programmation empirique de recherche soucieuse de jeter un éclairage rigoureux sur des pratiques, objets et terrains numériques nais-

sants. Notre posture compréhensive s'inscrit plutôt en réponse à la popularité croissante des recherches numériques évaluatives ou prédictives dont l'échelle et la portée, quoique attrayantes et potentiellement génératrices de découvertes, tendent à réduire la réalité somme toute complexe des sujets étudiés et à la dépouiller de ses nuances.

En somme, cet ouvrage a l'ambition de proposer non seulement un état de l'art, mais également un point de vue original sur la recherche en contexte numérique dans le but de poursuivre la discussion au-delà de nos cercles. Il s'adresse donc tant aux personnes qui mènent des recherches empiriques en sciences humaines et sociales qu'à celles qui enseignent la méthodologie ou qui souhaitent affiner leurs connaissances et leur compréhension de la méthodologie au prisme du numérique.

Le plan de l'ouvrage

La première des quatre sections du livre, « Les méthodes en contexte numérique », s'intéresse à l'incidence du numérique sur les méthodes ainsi que sur la recherche comme telle, et vice versa. Cette section débute avec un chapitre de Florence Millerand, David Myles et Serge Proulx qui offre une réflexion critique sur la classification des différentes méthodes de recherche en contexte numérique et sur leurs implications épistémologiques. Pour ce faire, ils reprennent une cartographie proposée par Noortje Marres qui, dans le cadre de la recherche en contexte numérique, distingue quatre approches méthodologiques (traditionnelle, computationnelle, virtuelle et numérique). Dans le deuxième chapitre, Dominique Boullier s'interroge sur la manière dont les sciences sociales exploitent les traces massives fournies par les médias socionumériques à l'aide d'un questionnement sur la redistribution des « agentivités ». Il postule que la traçabilité généralisée fournit une nouvelle métrique fondée sur les propagations, qui permet aux sciences humaines et sociales d'adopter un point de vue inédit sur les phénomènes de société. Le chapitre d'Alexandre Coutant et de Jean-Claude Domenget propose quant à lui un guide méthodologique pour analyser le numérique dans une perspective complexe mobilisant la sociologie des usages. Prenant pour illustration une recherche menée sur les refus de prescription algorithmique, les auteurs fournissent des pistes concrètes pour s'assurer de la cohérence entre les ambitions scientifiques affichées en sociologie des usages et la conduite des enquêtes empiriques.

La deuxième partie, « Les méthodes ethnographiques », est consacrée aux démarches ethnographiques ou ayant une composante ethnographique. Cette section s'ouvre avec la traduction inédite d'un texte de Christine Hine, figure incontournable de l'ethnographie en ligne. Le chapitre aborde les différents enjeux qui se présentent à l'ethnographe souhaitant étudier les communautés en ligne, ainsi que les pratiques ou les aspects culturels qui prennent forme ou sont reproduits en contexte numérique. Le chapitre suivant met en relation l'ethnographie et l'analyse du discours dans les recherches en communication. David Myles y montre les tensions entre ces méthodes, mais aussi le potentiel qu'elles représentent lorsqu'on les combine afin d'appréhender de manière complexe un phénomène en ligne. Sandrine Roginsky défend quant à elle la triangulation méthodologique comme principe pour étudier l'usage des médias sociaux en contexte politique, à la fois en ligne et hors-ligne. Ce chapitre présente en détail la stratégie employée dans le cadre d'une enquête longitudinale et montre comment articuler un bricolage méthodologique de manière cohérente. Dans la même veine, Christine Thoër, Florence Millerand et Nina Duque présentent un assemblage de méthodes, à la fois traditionnelles et numériques, afin de cerner le phénomène du visionnement connecté chez les jeunes. Elles montrent comment la combinaison des méthodes permet d'aborder différentes facettes du visionnement et de pallier les limites inhérentes à chacune. Le chapitre de Nina Duque présente les enjeux méthodologiques d'une ethnographie visant à cerner la « chambre adolescente ». L'autrice y soulève les défis de la recherche auprès des jeunes et la complexité inhérente à la prise en compte de multiples sites numériques du monde adolescent, mais également leur imbrication dans les dimensions physiques de la chambre. Enfin, le texte de Claire Balley aborde les représentations de l'intimité adolescente dans des vidéos publiées sur YouTube. Elle y détaille le processus de constitution de la problématique et de circonscription du terrain en illustrant les choix réalisés dans le cadre d'une enquête ethnographique en ligne sur la mise en scène de soi des jeunes.

La troisième partie aborde « Les méthodes sur traces », soit celles qui tirent parti des potentialités des traces produites dans les médias sociaux numériques par les usages qu'en font les internautes (c'est-à-dire toute trace laissée volontairement ou non dans le cadre de pratiques, comme des publications originales, des commentaires, des mentions « j'aime », le

partage de photos et de vidéos, mais aussi le fait que chacune de ces actions génère des métadonnées, etc.). Cette section s’amorce avec un chapitre de Guillaume Latzko-Toth, Claudine Bonneau et Mélanie Millette, qui réitérent la pertinence des recherches qualitatives sur de petits corpus, considérant ce type de démarche comme alternative et complémentaire aux travaux sur les données massives. Ils développent un modèle de densification des données et détaillent sa structure en trois couches afin qu’elle soit adaptable à différents types d’enquête sur des objets et des terrains numériques. Dans le chapitre suivant, Nicole Gallant, Katherine Labrecque, Guillaume Latzko-Toth et Madeleine Pastinelli décrivent une méthode d’entrevue dite « sur traces », qui consiste à faire commenter aux internautes les traces de leurs usages. La visite commentée permet de documenter les contextes de production de ces traces numériques dans le but de limiter leur sur- ou sous-interprétation par l’analyste. Le chapitre suivant, signé par Daniel Pelissier, présente une technique d’entretien augmentée d’un dispositif vidéo ayant l’avantage de documenter les usages qui ne laissent pas de traces visibles sur les plateformes. L’auteur aborde la manière dont la méthode de l’autoconfrontation d’une personne à sa propre trajectoire de navigation sur un site Web permet d’expliquer les expériences personnelles de manière à révéler la dimension active de pratiques habituellement invisibles. Enfin, le chapitre de Claudine Bonneau aborde les défis liés à l’analyse de corpus en ligne qui ne sont pas structurés d’avance (à la différence des groupes Facebook ou des ensembles de publications agrégées autour d’un mot-clic sur Twitter, par exemple). La chercheuse relève les avantages de la collecte manuelle pour contourner ces défis dans le cadre d’une recherche sur les pratiques de travailleurs et travailleuses visant à rendre visible la face cachée de leur réalité professionnelle.

La quatrième et dernière partie de l’ouvrage, « Les méthodes d’analyse d’images et les dispositifs vidéo », fait écho à une tendance de plus en plus marquée dans les médias numériques et les technologies numériques de communication. Cette tendance se traduit à la fois par la prégnance de l’image dans les corpus et les pratiques à l’étude, et par la place grandissante des méthodes qui s’intéressent aux images numériques, fixes ou en mouvement, ainsi qu’à leur trajectoire et à leur ordonnancement selon l’usage de métadonnées ou d’algorithmes. Les chapitres de cette partie tirent profit des innovations techniques pour proposer des stratégies méthodologiques originales, en commençant par celui de Nathalie

Casemajor, qui traite des métadonnées des images en ligne et de leur exploitation dans le cadre d'enquêtes qualitatives et ethnographiques. L'autrice y détaille les rouages des métadonnées et les informations qu'elles recèlent afin de montrer leur contribution à l'analyse visuelle du Web. Dans le chapitre suivant, Julien Figeac et Johann Chaulet présentent un dispositif d'enquête avec double captation vidéo qui permet de documenter de manière inédite les usages de la téléphonie mobile. Leur méthodologie appelle à prendre en compte à la fois le parcours écran de l'appareil mobile et le contexte d'usage. Le dernier chapitre de l'ouvrage, signé par Gabrielle Silva Mota Drumond, présente une enquête sur la représentation des usages qui est configurée à même les algorithmes de la plateforme de vidéo en ligne Netflix. L'autrice montre comment un assemblage de méthodes faisant une large place à l'analyse de la grammaire visuelle des interfaces permet de rendre plus transparents les processus et les implications des algorithmes de recommandation qui organisent les contenus de la plateforme. Finalement, une postface signée par Serge Proulx résume la contribution de l'ensemble des chapitres de l'ouvrage et en discute les implications pour la recherche en sciences humaines et sociales. L'auteur avance que la pertinence des recherches qualitatives mobilisant le numérique à la fois comme outil et objet d'enquête sociale se trouve renouvelée alors que les modélisations quantitatives suscitent de plus en plus d'attention médiatique et politique.

Les méthodes qualitatives et mixtes demeurent incontournables en sciences humaines et sociales pour rendre compte du caractère situé des pratiques numériques et offrir une réflexion nuancée sur leurs implications culturelles, économiques et sociopolitiques. Cet objectif est d'autant plus important dans le contexte d'une ruée vers les données massives, d'une privatisation galopante des espaces en ligne et de la prédominance des données « propriétaires » qui en sont issues, et qui y sont produites de manière à rendre certaines pratiques sociales plus visibles, intelligibles et légitimes que d'autres. Ces jeux de visibilité différenciée s'avèrent difficilement saisissables par les méthodes computationnelles automatisées. Les approches qualitatives proposées dans cet ouvrage s'inscrivent en rupture avec ces rapports asymétriques dans la production de traces numériques en misant sur des stratégies ingénieuses pour rendre compte des mutations liées au numérique, à son étude sociale et à son appropriation méthodologique par les chercheuses et les chercheurs.

Références

- ANANNY, Mike et Kate CRAWFORD, « Seeing without knowing: Limitations of the transparency ideal and its application to algorithmic accountability », *New Media & Society*, vol. 20, n° 3, 2016, p. 973-989.
- BARATS, Christine (dir.), *Manuel d'analyse du Web en sciences humaines et sociales*, Armand Colin, 2013.
- BAYA-LAFFITE, Nicolas et Bilel BENBOUZID, « Présentation: Imaginer la sociologie numérique », *Sociologie et sociétés*, vol. 49, n° 2, 2017, p. 532.
- BEUSCART, Jean-Samuel, Éric DAGIRAL et Sylvain PARASIE, *Sociologie d'Internet*, Armand Colin, 2016.
- JOUËT, Josiane, « Retour critique sur la sociologie des usages », *Réseaux*, n° 100, 2000 p. 487-521.
- LATZKO-TOTH, Guillaume, Claudine BONNEAU et Mélanie MILLETTE, « Small data, thick data: Thickening strategies for trace-based social media research », dans Anabel QUAN-HAASE et Luke SLOAN (dir.), *The SAGE Handbook of Social Media Research Methods*, SAGE Publications, 2017, p. 199-214.
- LUKA, Mary Elizabeth et Mélanie MILLETTE, « (Re)framing big data: Activating situated knowledges and a feminist ethics of care in social media research », *Social Media + Society*, vol. 4, n° 2, 2018.
- MILLERAND, Florence, Serge PROULX et Julien RUEFF (dir.), *Web social. Mutation de la communication*, Presses de l'Université du Québec, 2010.
- PROULX, Serge, *Penser les usages des technologies de l'information et de la communication aujourd'hui: enjeux, modèles, tendances*, Conférence inaugurale du colloque « Enjeux et usages des TIC. Aspects sociaux et culturels », Université de Bordeaux III, 22 septembre 2005.

PREMIÈRE PARTIE

Les méthodes en contexte numérique

CHAPITRE 1

La redistribution des méthodes de recherche en contexte numérique : critique d'une cartographie

Florence Millerand, David Myles, Serge Proulx

L'accès à un vaste éventail d'activités sociales en ligne offre de nouvelles occasions de recherche et des données originales pour appréhender des phénomènes sociaux inédits ou méconnus. D'une part, l'environnement numérique ouvre de nouvelles voies d'accès aux terrains de recherche et à des jeux de données de plus en plus nombreux provenant d'institutions publiques et de l'industrie privée; d'autre part, il renouvelle les objets d'étude en permettant d'observer des pans entiers de l'expérience des individus sur le Web, parfois *in vivo*, par exemple sur les médias sociaux, et possiblement sur une plus grande échelle. Les promesses liées aux données massives ne soulèvent pas l'enthousiasme des seuls *data scientists*, mais suscitent aussi l'intérêt des sciences sociales interpellées par les traces laissées par les actions quotidiennes sur le Web. La diversité des approches méthodologiques, mais aussi les réaffirmations fortes qu'elles soulèvent – quant à la place de la théorie, au statut des données, au rôle des scientifiques, à l'importance du contexte (Proulx et Rueff, 2018) – interrogent les conditions de réalisation de la recherche « sur » et « avec » le numérique.

On pourrait être tenté de penser le renouvellement des méthodes pour la recherche en contexte numérique comme étant fondamentalement nouveau, rompant avec les méthodes existantes. Selon Noortje Marres (2012),

la reconfiguration du paysage méthodologique en cours s'apparente plutôt à une « redistribution » des méthodes, et cette redistribution serait en partie le fruit de leur adaptation aux caractéristiques du numérique. Marres avance que cette redistribution participe plus largement d'une reconfiguration des relations entre les personnes impliquées dans les activités de recherche aboutissant, entre autres, à des redistributions de compétences, de pouvoir et de légitimité. Ces reconfigurations posent des défis épistémologiques importants et offrent en même temps l'occasion de revisiter les présupposés des méthodes de recherche en sciences sociales.

Ce chapitre propose une critique réflexive de l'hypothèse d'une redistribution des méthodes suggérée par Noortje Marres pour penser les conséquences – ou l'absence de conséquences – du virage numérique sur les méthodes en sciences sociales. La typologie proposée par Marres permet d'esquisser une cartographie des méthodes et de présenter une vue d'ensemble des grandes orientations méthodologiques qui structurent les recherches en contexte numérique. Il importe d'examiner l'hypothèse de la redistribution tant des méthodes que des acteurs de la recherche à partir des possibilités que cette proposition ouvre pour réfléchir à la construction sociale de la connaissance au temps de la transition numérique, ainsi que de réfléchir à l'application du concept de redistribution aux dispositifs eux-mêmes.

Une cartographie des méthodes de recherche en contexte numérique

L'idée selon laquelle les technologies numériques ont contribué à modifier les processus de recherche en sciences humaines et sociales semble faire consensus depuis le « tournant computationnel » des années 1980 (Berry, 2011). Les opinions varient cependant en ce qui concerne la nature et la portée de ces transformations sur le plan des méthodologies. Certains entrevoient un potentiel important d'innovation méthodologique – avec des implications ontologiques et épistémologiques (Boullier, 2015; Rogers, 2013), tandis que d'autres envisagent ces transformations dans la continuité du développement des méthodes surtout quantitatives (boyd et Crawford, 2011). Ces débats, quoique pertinents, entretiennent une certaine confusion dans l'évaluation de la portée des changements en cours, d'autant plus qu'il est parfois difficile de se repérer dans la diversité des méthodes préconisées dans les recherches.

Un premier apport notable du travail de Marres consiste à proposer une vue d'ensemble sur les grandes orientations méthodologiques structurant les recherches. Marres distingue quatre grands ensembles de méthodes : (1) les méthodes traditionnelles (ou *methods-as-usual*), (2) les méthodes computationnelles (ou *big methods*), (3) les méthodes virtuelles (ou *virtual methods*) et (4) les méthodes numériques (ou *digital methods*). Nous reprenons ici ces quatre grandes catégories pour esquisser une cartographie des principales démarches méthodologiques en les situant dans les débats sur le potentiel d'innovation méthodologique associé au numérique.

Les méthodes traditionnelles

Cette première catégorie renvoie à l'utilisation de méthodes classiques des sciences humaines et sociales pour étudier les phénomènes numériques sans chercher à les adapter. Il s'agit de transposer les démarches et façons de faire habituelles à l'environnement numérique. La recherche sur le visionnement connecté présentée dans le chapitre de Thoër *et al.* (le présent ouvrage) en constitue un bon exemple, en utilisant des techniques de recueil de données traditionnelles (entrevues individuelles, groupes de discussion, carnets de temps, enquêtes par questionnaire) pour cerner l'inscription des pratiques de visionnement de contenus audiovisuels en ligne dans les sociabilités juvéniles.

Les raisons motivant ce type de démarche varient considérablement. Certains travaux justifient cette posture simplement par la pertinence de la stratégie et des méthodes mobilisées au regard du questionnement de la recherche. D'autres sont guidés par une vision conservatrice, voire « sceptique quant à la possibilité que les méthodes sociales subissent une transformation significative dans des environnements numériques » (traduction libre de Marres, 2012, p. 149) allant jusqu'à remettre en question les fondements du tournant computationnel ou soulignant les dangers qui y sont associés. Les travaux utilisant de manière volontaire des méthodes traditionnelles peuvent ainsi chercher à se distancier d'un discours sur le potentiel transformateur du numérique dans une perspective se voulant critique. Cela étant dit, en ce qui concerne les recherches quantitatives en particulier, peu contestent le fait que de très grands corpus de données numériques diversifiées et hétérogènes montrent rapidement les limites

des techniques d'analyse traditionnelles fondées sur l'agrégation et la corrélation de données relativement simples avec un nombre limité de variables classiques (Venturini *et al.*, 2014), obligeant ainsi à faire appel à de nouvelles méthodes hybrides de traitement et d'analyse, fondées notamment sur l'apprentissage automatique (*machine learning*). Les nouvelles possibilités en matière de traitement automatisé de grands volumes de données sont au cœur des méthodes dites computationnelles.

Les méthodes computationnelles

Les méthodes computationnelles (*big methods*) sont associées aux outils issus des domaines de l'informatique et des mathématiques pour exploiter « les opportunités que les médias en ligne offrent pour le développement de réseaux à grande échelle et l'analyse textuelle [automatisée] » (Marres, 2012, p. 149). Ces techniques sont communément liées à l'extraction, la gestion, l'analyse et la visualisation de vastes corpus de données (*big data*) notamment – mais pas exclusivement – issues du Web et des médias sociaux. Les résultats de ces recherches quantitatives prennent généralement la forme de vastes réseaux structurels ou de représentations visuelles complexes qui visent à capter et à reproduire des grands patrons de comportements, d'opinions et de relations. À titre d'exemple, les travaux du « Politoscope » (Paris) proposent une « méthodologie intégrée pour la collecte de données, la reconstruction, l'analyse et la visualisation du développement du paysage politique d'un pays à partir des données de Twitter » (Gaumont *et al.*, 2018, p. 1).

Contrairement à la première posture, celle-ci considère que le numérique a rendu possible de nouveaux développements dans les techniques de modélisation et, donc, de nouvelles capacités de prédiction, grâce à l'analyse de dynamiques de réseaux plus en prise avec le « monde réel » et sur des échelles auparavant rarement possibles (Marres, 2012, p. 149). Ces méthodes sont directement liées au fait que des plateformes (moteurs de recherche et médias sociaux), mais aussi des fournisseurs d'Internet et des acteurs de l'industrie privée, ont constitué de très grandes bases de données qui fournissent l'infrastructure nécessaire pour une science des réseaux à grande échelle.

Comme l'indiquent Latzko-Toth *et al.* (le présent ouvrage), de nombreux travaux du champ des *Internet studies* ont souligné les limites

associées aux recherches mobilisant les méthodes computationnelles pour l'analyse des traces numériques laissées par les internautes. Ces travaux soulignent, entre autres, la tendance de ces recherches à reposer sur une logique strictement positiviste prônant la possibilité d'accéder à une réalité sociale unique et tangible et à considérer les données recueillies comme étant *de facto* naturelles et neutres (boyd et Crawford, 2011). Ces limites reposent sur deux confusions associant quantité et représentativité d'une part, et automatisation et objectivité, d'autre part. À la croyance selon laquelle « en raison de leur volume exorbitant, les données numériques auraient suffisamment "absorbé" le réel pour le représenter plus efficacement que toutes autres techniques d'objectivation » (Venturini *et al.*, 2014, p. 10) se superpose celle d'une « neutralité des calculs parvenant, sans hypothèses préalables, à produire des résultats objectifs » (*ibid.*, p. 14). En réponse à ces limites, un nombre croissant de recherches appellent à considérer la nature située, complexe et construite des traces recueillies en ligne, prônant ainsi une logique de données denses (*thick data*) – en opposition ou, du moins, en complémentarité à la logique des données massives (*big data*) – reposant sur une description analytique dense et contextuelle des traces et des usages numériques.

Les méthodes virtuelles

Les méthodes virtuelles renvoient aux recherches mobilisant une perspective ethnographique pour étudier des phénomènes socioculturels en contexte numérique, comme un forum de discussion de personnes vivant avec une même maladie ou une plateforme de jeu en ligne. Ancrées dans la tradition socioanthropologique, les méthodes virtuelles ont été grandement popularisées par les travaux de Christine Hine depuis le début des années 2000 (voir Hine, le présent ouvrage). Elles se déclinent sous différentes formes : ethnographie virtuelle ou du Web (Hine, 2015), netnographie ou cyberethnographie. De manière générale, ce type d'approche repose sur l'adaptation de méthodes qualitatives classiques de sciences sociales, principalement ethnographiques, à l'environnement numérique. Comme le souligne Marres (traduction libre, 2012, p. 150), ce programme de recherche met l'accent

sur les possibilités offertes par la transposition des méthodes de recherche sociale qualitative dans des environnements numériques en ligne. Sa principale

préoccupation [est] la transformation numérique de nos propres méthodes sociologiques, c'est-à-dire la façon dont des méthodes telles que l'analyse du discours et l'ethnographie ont été et pourraient être transformées par leur application dans un nouveau contexte.

Concrètement, il s'agit d'adapter des outils traditionnellement utilisés dans la recherche qualitative, comme l'observation (qu'elle soit participante ou non), l'entrevue (*in situ*, à distance, sur traces) et les analyses textuelles manuelles (analyse du discours, analyse thématique, analyse de conversation, analyse des interactions, analyse sémiotique) pour les utiliser dans l'environnement numérique (voir Myles, le présent ouvrage).

Au cœur de ces approches, nous retrouvons le principe ethnographique selon lequel le chercheur ou la chercheuse fait partie intégrante du dispositif de recherche et participe à part entière à la production des matériaux servant de base à l'analyse. L'un des avantages des approches ethnographiques en ligne réside dans cette capacité à étudier de façon « détaillée et approfondie [l]es pratiques sociales et [le] sens que leur prêtent les individus » (Hine, le présent ouvrage). Cette attention portée à l'inscription des activités dans leur contexte « a permis de démontrer la nature sociale complexe des interactions sur Internet et d'explorer les nouvelles formations culturelles émergeant en ligne » (*ibid.*).

Les méthodes virtuelles se sont considérablement transformées en cherchant à suivre les changements des dispositifs et de leurs usages. Trois grandes phases se sont ainsi succédé (Robinson et Schulz, 2009). Les premiers travaux ethnographiques étaient centrés sur les questions identitaires et visaient à montrer le caractère distinct des formes sociales en ligne. Le deuxième ensemble de travaux a cherché à se distancier de cette posture en soulignant plutôt la pertinence de transposer les questionnements méthodologiques hors-ligne à l'environnement en ligne, comme l'intégration de récits autoethnographiques pour mieux comprendre des formes de sociabilité en ligne, par exemple. Plus récemment, les travaux ont cherché à contextualiser les phénomènes observés dans les espaces où ils émergent, dans le cadre d'approches multimodales intégrant les données vidéo et audio au côté des données textuelles traditionnelles. Christine Hine résume ces déplacements en évoquant un recentrage des réflexions sur les manières de faire de l'ethnographie pour Internet plutôt que d'Internet ou encore par Internet.

Les critiques formulées à l'endroit des méthodes virtuelles sont venues principalement du développement des méthodes numériques amorcé

notamment par Richard Rogers (2009) dans sa conférence inaugurale au titre évocateur, *The End of the Virtual: Digital Methods*. Le reproche majeur porte sur le fait que les méthodes virtuelles ne reconnaîtraient pas pleinement la spécificité de l'univers numérique en considérant que l'importation des méthodes classiques des sciences humaines et sociales, même adaptées, suffit à l'étude des phénomènes numériques. Or, les méthodes numériques plaident justement pour le développement de méthodes « nativement numériques ».

Les méthodes numériques

Comme les méthodes virtuelles, les méthodes numériques (*digital methods*; Rogers, 2009) cherchent à appréhender des phénomènes socio-culturels tels qu'ils se manifestent en ligne, mais en mobilisant des outils issus directement de l'environnement numérique – par exemple, la comparaison entre des recherches réalisées dans différents lieux géographiques avec les mêmes moteurs de recherche ou les statistiques fournies par Google Analytics – et en les mettant au service d'objectifs scientifiques. Les démarches sont souvent mixtes et tablent sur la complémentarité de méthodes virtuelles et computationnelles. À ce titre, elles se distinguent des méthodes computationnelles par l'utilisation de stratégies à la fois quantitatives et qualitatives. Les travaux menés sur la constitution de publics *ad hoc* (sans liens préexistants) autour de sujets particuliers sur Twitter, en se servant du mot-clic (*hashtag*), et mobilisant des méthodes mixtes (Bruns et Burgess, 2015) constituent de bons exemples de recherche s'inscrivant dans ce courant.

Comme le souligne Marres, ce courant méthodologique cherche davantage à adapter, contourner, bricoler ou réaffecter (*repurpose*) les outils mis à disposition des internautes ordinaires pour en faire des outils et méthodes de recherche. Ainsi, le programme des méthodes numériques repose avant tout sur la volonté de « tirer profit des capacités analytiques et empiriques qui sont incorporées » aux technologies numériques du quotidien (traduction libre de Marres, 2012, p. 151) pour former des méthodes inédites qualifiées de « nativement numériques » (*natively digital*; Rogers, 2013). Dans cette perspective, les dispositifs constitutifs d'Internet, comme l'hyperlien, le site Web ou les moteurs de recherche, sont considérés comme étant les plus aptes à pouvoir informer les recherches en participant

à part entière à la production de catégories émiqes permettant d'interpréter les pratiques en ligne. La démarche vise à tirer du Web lui-même les outils pour l'étudier en les adaptant éventuellement afin d'aboutir à des dispositifs d'enquête et à des méthodes « moulés » au numérique pour en préserver toutes les propriétés. Rogers propose l'expression « *follow the medium* » pour qualifier ce principe méthodologique visant à appréhender le numérique à la fois comme objet d'étude et instrument d'analyse.

Un avantage important des méthodes numériques est de produire des méthodes sur mesure et adaptées au contexte des phénomènes étudiés. Cependant, elles s'interrogent sur la nécessité de créer des méthodes particulières pour appréhender les objets numériques et posent la question de la pérennité des recherches et des connaissances produites. En effet, ces méthodes sont fortement tributaires de l'instabilité qui caractérise l'environnement numérique. Le fait que les dispositifs changent constamment à la fois dans leurs fonctionnalités et leurs politiques d'accès contribue au caractère fortement situationnel et expérimental de ce type de méthodes.

Enfin, une autre limite importante des méthodes numériques concerne le risque de faire de la science à partir de « boîtes noires », dans la mesure où les recherches peuvent manquer de réflexivité face aux logiques internes des dispositifs numériques lorsque celles-ci sont acceptées sans discernement. Comme le souligne Marres, le fait que les politiques d'organisations privées et d'entreprises déterminent les conditions d'accès des scientifiques aux données représente un inconvénient majeur. Pour éviter le risque d'aveuglement inhérent à ces méthodes, Plantin et Monnoyer-Smith (2013) invitent à un nécessaire effort de déconstruction réflexive des catégories de l'analyse, faisant preuve d'une prudence méthodologique en résonance avec le principe épistémologique énoncé jadis par Gaston Bachelard de la nécessaire « coupure épistémologique » de la démarche scientifique vis-à-vis des catégories du sens commun. D'un côté, il importe donc de militer pour avoir un accès aux données – sans lequel les méthodes numériques ne peuvent se déployer entièrement – afin de procéder à une recherche indépendante et de qualité. De l'autre, il est nécessaire de faire montre d'une certaine prudence face à ces méthodes et de remettre en question les catégories imposées par l'industrie du numérique.

Limites de la cartographie proposée par Noortje Marres

Si la cartographie suggérée par Noortje Marres propose un point de vue intéressant sur la reconfiguration du paysage méthodologique en sciences sociales actuellement en cours, elle présente également certaines limites. D'abord, cette cartographie tend à mettre sur un même plan des approches qui se distinguent par des positionnements très différents dans le champ scientifique. Ainsi, les méthodes computationnelles, fondées sur des techniques de traitement de données volumineuses provenant à l'origine du monde de l'informatique, sont utilisées désormais par un large éventail de champs disciplinaires. Les méthodes virtuelles et les méthodes numériques, en revanche, restent pour le moment cantonnées aux sciences humaines et sociales, voire au champ des *Internet studies*.

Une deuxième limite de cette cartographie concerne sa tendance à reproduire les clivages traditionnellement établis entre les approches qualitatives et quantitatives alors que de nombreuses recherches prônent le développement de méthodes mixtes misant sur la complémentarité des approches pour aborder les phénomènes numériques. Si les analyses issues de méthodes virtuelles ont le potentiel d'éclairer des résultats de recherches quantitatives, les analyses quantitatives des traces numériques « peuvent offrir davantage de quantification, mais seulement au prix d'une meilleure qualification » (Venturini *et al.*, 2014, p. 16), ce qui exige de les resituer dans leur contexte de production.

Enfin, cette représentation des approches tend à favoriser le développement et l'adoption des méthodes numériques (*digital methods*) – certes de manière critique, comme le suggère Marres – mais au détriment des autres types de méthodes, et notamment des méthodes virtuelles, jugées trop limitées pour pouvoir appréhender l'ampleur des changements suscités par le numérique lui-même sur les phénomènes sociaux. L'ordonnement des quatre catégories de méthodes en fonction d'une forme de redistribution minimale de la capacité de recherche (méthodes traditionnelles) à une redistribution maximale (méthodes numériques) sous-tend une évolution naturelle des outils de collecte et d'analyse en contexte numérique, omettant de considérer qu'une sélection opportune des méthodes s'effectue plutôt en fonction des objectifs de recherche visés.

Une redistribution des méthodes et des acteurs de la recherche

La proposition de Marres sur la redistribution des méthodes doit être comprise dans le contexte plus large des débats autour des enjeux du numérique pour la recherche en sciences humaines et sociales. Son principal intérêt réside, au-delà de la cartographie des méthodes qu'elle propose, dans les possibilités qu'elle offre pour examiner le problème de la construction sociale de la connaissance. L'argument de la « redistribution » implique que l'activité de recherche en contexte numérique s'apparente à un « accomplissement distribué » dans le sens où « les plateformes, les usagers, les dispositifs et [les] pratiques informationnelles contribuent activement à la réalisation de la recherche » (traduction libre de Marres, 2012, p. 139). Le caractère distribué de la recherche serait encore plus visible et « difficile à nier » dans le cas des recherches « sur » et « avec » le numérique (*ibid.*, p. 160), dans la mesure où les recherches universitaires intègrent de plus en plus de pratiques, de dispositifs et de données qui en sont directement issus. Il s'agit, par exemple, d'utiliser des informations produites par les dispositifs d'analyse (*analytics*) des plateformes en ligne (nombre de liens, nombre d'abonnements, etc.) comme des sources de données ou comme des outils de visualisation. Marres propose de considérer ces distributions non pas comme de nouveaux problèmes posés aux milieux scientifiques préoccupés par le numérique, mais comme une occasion de penser le réaménagement des méthodes de recherche en sciences sociales. C'est sur le terrain des méthodes que cette redistribution serait la plus tangible et la plus apte à mener à des innovations méthodologiques.

L'élargissement de la recherche à de nouveaux acteurs

Une première lecture de ces processus de redistribution met en évidence un élargissement important des types d'acteurs impliqués dans l'analyse de la vie sociale, voire une occasion pour la recherche « amateur » en sciences sociales et humaines de se déployer sur ce nouveau terrain d'études. La prolifération des données mises à disposition sur le Web, mais aussi des dispositifs pour les analyser favoriserait la « délocalisation » de la recherche et pourrait faire de « l'analyse sociale une pratique dans laquelle “n'importe qui pourrait vouloir s'engager” » (traduction libre de Marres, 2012, p. 142). Certaines réflexions récentes invitent ainsi à faire

de l'exploration des données numériques un moyen de dialogue avec les publics en ligne (Kennedy et Moss, 2015).

Une lecture plus critique souligne plutôt dans ce déplacement deux tendances préoccupantes. La première est le risque d'une concentration des ressources de la recherche dans quelques centres financièrement ou technologiquement disposés à négocier des ententes avec les plateformes productrices de données, par exemple au sein de grandes entreprises informatiques, ce qui se traduirait par une perte d'influence de la recherche publique au profit de l'industrie privée. La deuxième est liée à la montée en puissance des plateformes productrices de données qui contrôlent non seulement l'accès aux données d'usage, mais aussi leur organisation et mise en forme, imposant de fortes contraintes d'accès à la fois techniques et légales.

L'invitation à se pencher sur les processus de redistribution en cours vise à attirer l'attention sur les conséquences liées à la prolifération des plateformes et des outils permettant d'observer la vie sociale en ligne et, plus précisément, sur les processus de « reconfiguration des relations » entre les différentes catégories d'acteurs participant à la production des connaissances sur les phénomènes numériques (traduction libre de Marres, 2012, p. 144). La notion de redistribution provient de travaux menés dans le domaine des STS (*science and technology studies*), qui définissent la production de connaissances ou l'innovation technologique comme des interactions complexes entre un vaste ensemble d'acteurs situés à la fois à l'intérieur et à l'extérieur de l'université : sujets participants, agences de financement, infrastructures techniques, équipes de recherche, etc. Dans cette perspective, « la recherche est définie comme une entreprise collective impliquant un ensemble d'acteurs divers avec des rôles variés » (*ibid.*, p. 145). L'hypothèse de la redistribution de la recherche invite donc à étudier les nouvelles formes de division du travail de recherche et d'analyse qui sont au cœur de ces reconfigurations et à examiner leurs conséquences sur la fabrication sociale de la connaissance.

Les conséquences de cette redistribution sur la construction sociale de la connaissance

La présence des méthodes computationnelles engendre une redistribution des rôles. Les techniques de recueil et d'analyse de données impliquent

une division du travail entre sujets participants, dispositifs de collecte et analyses qui donne préséance aux techniques (et à leur puissance calculatoire) au détriment des sujets, voire des chercheurs et chercheuses. D'une part, les données produites par les activités en ligne sont considérées comme étant de meilleure qualité, car plus « objectives » que celles émanant des sondages réalisés auprès des sujets, perçues comme étant « subjectives » et donc sujettes à erreur, même s'il pourrait s'agir là d'un biais positiviste. D'autre part, un rôle clé est explicitement reconnu au milieu non scientifique constitué autour des plateformes et de leurs dispositifs d'enregistrement et de captation de données.

En ce qui concerne les méthodes virtuelles, la redistribution est partielle, mais là aussi significative. À l'opposé des méthodes computationnelles, ces méthodes donnent préséance aux sujets participants ainsi qu'aux scientifiques. L'expérience des sujets étant au cœur de ces approches, celles-ci tendent à maximiser leur rôle dans l'activité interprétative. En outre, l'ethnographe en ligne reste un élément clé du dispositif d'enquête pour comprendre les pratiques et motivations derrière les usages. Ces approches reconnaissent aux sujets une capacité réflexive dans la mesure où ils s'engagent dans des activités semblables à celles des ethnographes, comme tenir un journal pour documenter ses propres pratiques. Une redistribution réussie des méthodes ne limiterait pas le rôle actif des sujets participants et celui des chercheurs et chercheuses.

Enfin, les méthodes numériques sont celles qui impliquent la redistribution la plus importante en donnant la préséance aux dispositifs nativement numériques. Ces derniers participent à une transformation des conditions de réalisation de la recherche de deux manières : d'une part, en incorporant des méthodes traditionnelles et en les modifiant ; d'autre part, en jouant un rôle potentiellement structurant dans la production des catégories d'analyse. Ainsi, lorsqu'une recherche mobilise les données générées par Twitter (par exemple par des requêtes à son API) pour repérer les individus les plus centraux dans une discussion, celle-ci reproduit en réalité une catégorie produite par Twitter correspondant aux personnes les plus actives dans un classement de contribution (Marres, 2012, p. 145). En outre, les nombreuses analyses de réseaux réalisées à partir des contributions sur Facebook apparaissent largement guidées par les outils et métriques de la plateforme elle-même, y compris les limites qu'elle impose. Autre exemple, dans le cadre d'une analyse de conversa-

tions menée sur YouTube, la constitution d'un corpus de vidéos apparaît comme une activité distribuée dans le sens où l'algorithme de recommandation de YouTube rend certains contenus – et donc, certaines conversations – plus visibles que d'autres. Ce sont ces modifications et ces enjeux qui intéressent tout particulièrement les tenants des méthodes numériques cherchant à comprendre le rôle du numérique dans la production des connaissances, non seulement pour mieux en réaffecter les outils au service d'objectifs de recherche, mais aussi pour pouvoir en contester éventuellement les conséquences sur les nouvelles divisions du travail scientifique et sur sa production.

La reconnaissance de nouvelles formes de délégation

Selon Marres et Gerlitz (2015, p. 33), «la recherche sur les médias sociaux pourrait mettre en cause les relations entre les participants à la recherche, dont celles entre sujets et objets de recherche, ainsi que les hiérarchies supposées des acteurs et actants en recherche sociale». En l'occurrence, nous avons mis en évidence tout au long de notre texte le fait qu'un ensemble élargi d'acteurs et d'entités sont impliqués dans les processus de recherche (plateformes, usagers et usagères, bases de données, algorithmes, entreprises informatiques, scientifiques). Dans ces nouvelles formes d'organisation de la recherche, des entreprises privées mobilisant des dispositifs algorithmiques jouent un rôle inédit en contribuant à la fois à la production et à la gestion des corpus de données servant à l'analyse.

Porter attention à la nature des redistributions en cours implique de considérer les formes de délégation qui sont impliquées. En fin de compte, l'argument de Marres invite à mettre en lumière les nouveaux rôles attribués aux acteurs et aux dispositifs et, en particulier, à expliquer les conséquences de ces délégations sur nos capacités de recherche et d'analyse. Reconnaître le rôle des dispositifs dans la production de la connaissance n'est pas un phénomène nouveau. À ce titre, les études sur les infrastructures informationnelles en ont établi l'importance depuis longtemps. En revanche, l'attention sur la redistribution des méthodes offre une occasion d'étudier précisément comment les dispositifs numériques participent à la production des catégories d'analyse du social tout en étant eux-mêmes façonnés par divers mondes sociaux. Ces dispositifs augmentés redessinent les frontières entre recherche publique et secteur privé, entre

scientifiques, sujets et outils. Il devient crucial d'examiner les dispositifs en eux-mêmes pour analyser les différents arrangements sociotechniques qui les caractérisent et comprendre comment ils participent à créer de nouveaux régimes de connaissance.

Références

- BERRY, David M., «The computational turn: Thinking about the digital humanities», *Culture Machine*, vol. 12, 2011, p. 1-22.
- BOULLIER, Dominique, «Les sciences sociales face aux traces du *big data*», *Revue française de science politique*, vol. 65, n° 5, 2015, p. 805-828.
- BOYD, danah et Kate CRAWFORD, «Six provocations for big data», *A Decade in Internet Time: Symposium on the Dynamics of the Internet and Society*, 2011, p. 1-17.
- BRUNS, Axel et Jean BURGESS, «Twitter hashtags from ad hoc to calculated publics», dans Nathan RAMBUKANA (dir.), *Hashtag Publics: The Power and Politics of Discursive Networks*, Peter Lang, 2015, p. 13-28.
- GAUMONT, Noé, Maziyar PANAHİ et David CHAVALARIAS, «Reconstruction of the socio-semantic dynamics of political activist Twitter networks – Method and application to the 2017 French presidential election», *PLOS ONE*, vol. 13, n° 9, 2018, e0201879.
- HINE, Christine, *Ethnography for the Internet: Embedded, Embodied and Everyday*, Bloomsbury Publishing, 2015.
- KENNEDY, Helen et Giles MOSS, «Known or knowing publics? Social media data mining and the question of public agency», *Big Data & Society*, vol. 2, n° 2, 2015, p. 1-11.
- MARRES, Noortje, «The redistribution of methods: On intervention in digital social research, broadly conceived», *The Sociological Review*, vol. 60, n° 1, 2012, p. 139-165.
- MARRES, Noortje et Carolin GERLITZ, «Interface methods: Renegotiating relations between digital research, STS and Sociology», *The Sociological Review*, vol. 64, n° 1, 2015, p. 21-46.
- PLANTIN, Jean-Christophe et Laurence MONNOYER-SMITH, «Ouvrir la boîte à outils de la recherche numérique. Trois cas de redistribution de méthodes», *tic&société*, vol. 7, n° 2, 2013.
- PROULX, Serge et Julien RUEFF, *Actualité des méthodes de recherche en sciences sociales sur les pratiques informationnelles*, Centre d'études sur les médias, 2018.
- ROBINSON, Laura et Jeremy SCHULZ, «New avenues for sociological inquiry: Evolving forms of ethnographic practice», *Sociology*, vol. 43, n° 4, 2009, p. 685-698.
- ROGERS, Richard, *Digital Methods*, The MIT Press, 2013.
- ROGERS, Richard, *The End of the Virtual: Digital Methods*, Amsterdam University Press, 2009.
- VENTURINI, Tommaso, Dominique CARDON et Jean-Philippe COINTET, «Présentation», *Réseaux*, vol. 6, n° 188 : «Méthodes digitales. Approches quali/quantitative des données numériques», 2014, p. 9-21.

CHAPITRE 2

Les traces numériques et le pouvoir d'agir des répliques

Dominique Boullier

Les sciences sociales peuvent-elles apprendre de l'apprentissage qui domine désormais la science des données (*data science*) par l'apprentissage automatique (*machine learning*) ? La mutation culturelle ne peut être ignorée, mais cela reste difficile de la saisir, de ne pas tout rejeter d'un bloc (dont les fausses promesses prédictives) ou de ne pas succomber à la tentation (« Google est trop fort, l'académie a perdu ! »). En fait, il nous paraît possible à la fois de se laisser affecter par la vague des données massives (*big data*) et de l'apprentissage automatique et de profiter de cette occasion pour mettre en évidence tout ce que les sciences sociales peuvent apporter à ces méthodes, à condition d'éclaircir ce qu'elles font et de simplifier le message. Nous proposons de montrer comment des « modèles conceptuels », au sens de Norman (1988), irriguent toutes les sciences sociales et comment une vision claire de ces divers « points de vue » équipés permet de spécifier les contributions effectives des sciences sociales dans l'exploitation des traces fournies désormais en masse par les réseaux numériques. Paradoxalement, l'informatique nous oblige à expliquer ce que les sciences sociales font dans leur « distribution d'agentivité », à la modéliser et ainsi à la rendre transférable, communicable, voire calculable.

Nous prendrons comme point de départ les propriétés reconnues des données massives pour mesurer les glissements méthodologiques auxquels nous faisons face. La qualité des données n'est plus du tout évaluée selon les critères auxquels la statistique nous avait habitués (l'exhaustivité et la

représentativité ont été remplacées par le volume et la variété, auxquels a été ajoutée la vélocité). Nous proposons de convertir cette vélocité en qualité statistique pérenne, qui serait la traçabilité, de façon à exploiter des gisements de données dans le temps et avec un grain suffisamment fin. Cette évolution d'apparence technique sera mise en perspective dans une histoire des méthodes de quantification, qui s'inspire des travaux d'Alain Desrosières, mais qui met en place trois générations différentes, toutes appuyées sur des états différents des techniques : les recensements, les sondages et désormais les traces massives. Pour éviter de penser à ces générations comme à des tables rases successives, alors qu'elles gardent chacune leur pertinence, nous montrons comment chacune de ces méthodes de quantification encapsule un certain point de vue sur le social. Loin de prétendre écraser les autres points de vue, celui des répliques rendues observables par les traces massives sur les réseaux numériques permet d'accéder à une dimension du social, celle des propagations d'entités qui traversent les esprits humains et les supports numériques. La cohérence des choix d'objets de recherche, de méthodes et d'entités agissantes constitue une condition de clarification des débats entre paradigmes au sein des sciences sociales, qui reposent en fait sur des points de vue et des méthodes différentes.

Les données massives et les 3V comme ersatz

Les critères de qualité des données massives sont souvent résumés par les « 3V » : volume, variété, vélocité. Ces 3V fournissent des équivalents, sous forme d'ersatz, des exigences classiques des sciences sociales. Notre approche n'est pas contradictoire avec l'ontologie de Kitchin et McArdle (2016), mais elle prétend être plus pragmatique et historique à la fois, plus proche des travaux de Mackenzie (2017) sur le travail des *machine learners*. La définition de la « nature » des données massives que tentent de donner Kitchin et McArdle est ici reformulée en tableau des pratiques des différentes expertises en quantification selon les époques, des institutions qui les financent et des modèles du social qu'elles incorporent pour en comprendre les limites et les potentialités. Ainsi, comme nous allons le voir, il n'est guère possible d'ajouter l'exhaustivité comme critère des données massives puisque les traces que l'on collecte en ligne n'ont pas de limites dans leur diversité (variété) et sont produites de façon dynamique (vélo-

cité), ce qui n'empêche pas le calcul grâce aux bases de données NoSQL et aux techniques d'apprentissage automatique.

Le volume et l'exhaustivité

Le volume correspond plutôt à l'exigence d'exhaustivité qui a été établie lors de ce que nous avons appelé la première génération de sciences sociales (Boullier, 2015; 2016; 2017), qui reposait sur les recensements, équipée notamment par les machines de Hollerith qui deviendront IBM. En visant l'exhaustivité, et en l'atteignant au prix d'un investissement de forme considérable (Thévenot, 1986), les États et leurs instituts de statistiques permettaient aussi aux sciences sociales d'accéder à cette entité improbable, une « société » (nationale, rappelons-le). Durkheim en fit la théorie et lui attribua une agentivité quasi toute-puissante, sans pour autant s'appuyer sur ces recensements, mais plutôt sur des registres divers qui avaient cependant la prétention eux aussi d'être exhaustifs. Le pouvoir que nous continuons à attribuer à la « société », cette agentivité spécifique de la structure sociale, doit ainsi beaucoup au dispositif qui permet de prouver qu'elle existe bien puisqu'on la dénombre, quand bien même il s'agit avant tout de la gouverner (Desrosières, 2014).

Pour les données massives, l'exhaustivité n'est que rarement atteinte, parce que les flux de nouvelles données sont récoltés en permanence et parce que, lorsqu'il s'agit du Web, cette notion n'a même plus de sens puisque ses limites sont inatteignables. Mais le volume est une condition de possibilité requise pour que l'apprentissage puisse se faire (Domingos, 2015)¹. L'exhaustivité est ainsi dégradée en volume, mais cela permet d'agréger en revanche des traces de tous types en masse, en dehors des conventions lourdes que suppose l'exhaustivité. L'acceptation de l'approximation est à la base même des méthodes de l'apprentissage automatique, qui peut valider statistiquement cette marge d'approximation et apprendre

1. L'apprentissage sera d'autant plus performant qu'il pourra réviser des corrélations trop simplistes obtenues avec des observations trop peu abondantes (phénomène de surapprentissage ou *overfitting*). Les jeux de données d'apprentissage sur lesquels sont entraînés les algorithmes, souvent avec supervision pour leur dicter ce qui est correct et ce qui ne l'est pas, doivent être différents des jeux de données d'exécution sur lesquels les algorithmes entraînés devront découvrir eux-mêmes des corrélations robustes et prédire des résultats pertinents pour le problème concerné. Cette contrainte, entre autres, explique le besoin de volume suffisant.

au fur et à mesure des nouveaux exemples ou des nouvelles sources de données. Pour les sciences sociales et pour faire convention (Eymard-Duvernay *et al.*, 2004), il conviendra de fixer un équivalent de ce volume qui se rapproche des exigences traditionnelles de l'exhaustivité, sans pour autant pouvoir les respecter notamment lorsqu'on traite du Web. Nous devons clairement faire notre deuil de l'exhaustivité sans que cela nous dispense de fixer les cadres conventionnels de toute démarche en sciences sociales traitant de traces numériques. Lorsque l'on traite de faibles volumes avec de faibles capacités de calcul, il est quasiment inévitable de retomber sur des méthodes statistiques classiques comme les régressions linéaires, qui ne pourront pas traiter des problèmes d'une échelle de complexité trop élevée. La convention à développer devra donc obliger les travaux en sciences sociales/sciences des données à énoncer explicitement l'échelle de complexité du problème d'origine, les méthodes pour la réduire (à travers des réductions de dimensions ou variables, par exemple), la quantité d'observations nécessaires pour que le calcul soit significatif (tout en tenant des capacités de calcul disponibles), les autres méthodes abandonnées en cours de route et la justification de ces abandons. Toutes ces métadonnées sur un travail de science des données se trouvent parfois explicitées dans les challenges Kaggle (compétitions en science des données) que nous avons étudiés (Boullier et El Mhamdi, 2020) et qui montrent bien le caractère empirique des pratiques des *machine learners*.

La variété et la représentativité

Le deuxième critère, la variété, est lui aussi une forme de transcription de l'exigence de représentativité qui a permis à toutes les sciences sociales de procéder par enquêtes (sondages) à base d'échantillonnage. Cette innovation statistique date des années 1930 et sa mise en lumière en 1936 par Gallup (Blondiaux, 1998) a constitué un tournant dans la quantification des sociétés en faisant émerger une deuxième génération de sciences sociales. Toutefois, de nombreuses conditions ont dû être réunies pour que ces conventions se stabilisent et encadrent les pratiques des instituts de sondage qui les réalisent ainsi que celles des médias et des marques qui les financent. Mais ce montage spécifique ne peut être transféré tel quel dans le monde numérique, notamment lorsqu'il s'agit de données native-

ment numériques (Rogers, 2013). Il est donc irréaliste de prétendre mesurer « l'opinion » sur le Web sans être capable de respecter les critères de représentativité qui ont été établis depuis plusieurs décennies (avec les intervalles de confiance notamment) et les méthodes de recueil auprès d'individus d'expressions calibrées par des questionnaires. Dès lors, le critère de la variété doit être considéré comme une version lâche de la représentativité, qui suppose que l'on accepte un niveau « suffisant » de variété pour valider un échantillon.

Pour les sciences sociales de troisième génération, qui acceptent de perdre la contrainte de représentativité telle qu'elle a été construite dans le cas des sondages, il reste à définir ce que serait cette variété. La constitution d'un ensemble de sources suffisamment variées² devrait alors répondre à quelques critères propres aux méthodes numériques et au domaine étudié. Les travaux sur l'analyse de sentiment que j'ai publiés avec Lohard en 2012 nous ont conduits à nous a conduits à considérer qu'aucune description du social-société, du social-opinion ou du social-traces ne peut être produite en généralité sur les réseaux numériques. En effet, les particularités des plateformes et leur pouvoir de formatage des traces sont trop présents, comme le montre Rogers (2013), qui propose cependant d'effectuer une réorientation (« *repurposing* ») des traces à visée commerciale pour les recentrer sur les finalités de la recherche. La prolifération des traces rend paradoxalement impossible toute prétention à une référence à un « tout » posé *a priori* ou constitué *a posteriori*. Les sciences sociales doivent accepter de ne traiter que des enjeux (« *issues* ») (Marres, 2007; 2017; Marres et Weltevrede, 2013), ou des points de focalisation d'attention, dont le numérique peut garder les traces, des traces qui seront spécifiques à chaque enjeu. Cela réduit considérablement la portée totalisante des prétentions des données massives, mais cela rend possible une certaine forme de représentativité et d'exhaustivité.

La vélocité et la traçabilité

Enfin, le dernier critère, la vélocité, ne trouve guère d'équivalent dans les sciences sociales de première et deuxième génération. On parle de vélocité à deux points de vue : d'abord la vitesse de publication des données ou

2. On parle alors de méthodes de *sourcing* (« sources d'approvisionnement ») des sites Web, par exemple pour une étude de topologie comme dans l'étude que nous avons conduite sur le Web du livre en France à partir de plus de 5 000 sites (Le Béhec *et al.*, 2015).

traces, car elles arrivent en temps réel puisque chaque clic est immédiatement enregistré comme trace sur les serveurs des agences de placement publicitaire par exemple, ce qui produit un flux continu ; puis, la vitesse de leur traitement et de leur mise à jour, ce qui permet de réviser en permanence les calculs et les agrégations réalisées, faites parfois avec des approximations lorsque des données sont manquantes (ce qu'on ne pouvait pas faire sans des bases de données NoSQL), et surtout d'alimenter de façon dynamique et permanente l'apprentissage des algorithmes d'apprentissage automatique. À vrai dire, ces processus dynamiques ne constituaient ni le point fort, ni la préoccupation des sciences sociales. Il était en effet essentiel de trouver avant tout à représenter les positions à un instant donné, pour montrer la force d'imposition de « la société » sur la diversité des comportements individuels ou pour montrer comment l'opinion publique évoluait à partir des expressions singulières obtenues dans les enquêtes. Certes, grâce à un suivi longitudinal des mêmes populations ou la reprise des mêmes questionnaires, il était – et il est toujours – possible de restituer l'équivalent d'une dynamique, sans jamais cependant pouvoir suivre à la trace les médiations qui permettraient de produire ces évolutions. La vitesse semble donc hors du champ des approches plus classiques.

Cependant, une branche des sciences du Web s'est emparée de cette question de vitesse à sa façon, en exploitant les traces des mêmes qui se propagent sur le Web. Il est très significatif que Kleinberg, celui-là même qui avait exporté les méthodes de la scientométrie vers l'étude de la topologie du Web, méthodes qui furent reprises par Google, se soit intéressé depuis plusieurs années (2002) à la mise au point d'un « *meme tracker* » avec Leskovec (2009). Cette dernière étude, la plus connue, a porté sur la propagation des citations durant la campagne Obama, ce qui leur permit de réaliser une visualisation spectaculaire de la focalisation de l'attention en courbes à montées et descentes très rapides (*streams and cascades*) autour de certains incidents de la campagne. Leur méthode agrège tous les types de traces que peuvent laisser ces citations, traitées comme des chaînes de caractère dont on peut trouver la trace dans tout le Web, et en produit une métrique ancrée dans le temps, au jour le jour, voire minute par minute désormais avec Twitter (l'unité de mesure étant devenue le « *tweet per second* »). La prise en compte des mêmes nous paraît prometteuse sous réserve que l'on suive aussi les transformations-traductions ou

« dérivations » de ces mêmes dans des milieux différents, car « toute existence va différant », comme le disait Tarde (1895). Il sera alors possible de parler de mémétique (Dawkins, 1976 ; Dennett, 2017 ; Shifman, 2014) et non plus seulement de viralité.

Les conditions de félicité d'une traçabilité

Dès lors, il devient possible de trouver un équivalent de la vélocité des données massives : la traçabilité. Elle devient le critère essentiel de qualité des entités que l'on peut étudier. Quelques conditions de félicité doivent être réunies pour y parvenir :

- Les traces en question doivent avoir une continuité suffisante pour qu'il soit encore possible de dire qu'il s'agit d'un même processus, sans avoir la contrainte de l'étude de la viralité (propagation à l'identique) ni le laxisme d'une intertextualité généralisée, où tout pourrait être signe et reprise de tout.
- Les traces en question doivent permettre des suivis d'associations hétérogènes, c'est-à-dire une puissance de connectivité suffisante. Pour cette raison, des traces dont le format est trop spécifique à une plateforme peu connue ne peuvent donner lieu à extension ni à suivi, quand bien même il reste nécessaire d'expérimenter, dans un premier temps, les modèles de propagation des répliques sur une plateforme précise, plus aisément maîtrisable.
- Le suivi des traces en question doit permettre de dater avec précision tous les événements, toutes les transformations et toutes les associations. Les lignes de temps (*timelines*) sont ici l'équivalent d'autres conventions, comme les points cardinaux pour la topographie ou les inégalités de richesse pour les sciences sociales de première génération.

Ces conditions rendent possible le basculement des sciences sociales vers le suivi d'éléments qui ne sont plus ni des individus, ni des groupes, ni la société, ni l'opinion. Le numérique a dissous toute certitude sur leurs statuts, qui avaient été déjà remis en cause théoriquement, puisque les plateformes comme Facebook ou Google exploitent des attributs des comptes (614 de façon standard chez Facebook, sans compter ceux que les marchands de données y ajoutent [Andreou *et al.*, 2018]). Ces éléments peuvent être des comportements élémentaires dont on peut suivre la programmation (les « j'aime » ou les *retweets*), des mêmes issus de la sous-culture Internet (Reddit, 9gag par exemple), des citations reprises dans

plusieurs médias et plateformes, ou même des citations scientifiques entre revues, des slogans ou visuels publicitaires ou politiques, des expressions supposées individuelles et spontanées, mais qui sont en fait répétées par des millions de comptes en quelques heures, etc. La propagation de ces éléments devient l'objet d'étude de la troisième génération de sciences sociales, car ce sont les propriétés de ces entités qui leur permettent de créer de petites différences, et par là, de circuler et d'affecter les individus et les groupes, la société et l'opinion.

Cependant, ces conditions ne sont pas nécessairement considérées comme décisives pour obtenir un accord (et faire « convention ») de la part des plateformes (les GAFA en résumé) et des marques qui font vivre tout cet écosystème par l'exploitation publicitaire de cette traçabilité. De même que la sociologie durkheimienne s'est associée de fait avec les institutions étatiques productrices de données pour produire « la société » en combinant enquêtes (des statisticiens) et registres (des administrations) (Desrosières, 1993), les instituts de sondage de Gallup et de Lazarsfeld se sont associés aux médias grands consommateurs de données sur les publics pour produire leur « opinion publique », en rapprochant ainsi « mesures d'audience » (le public des médias) et « opinion publique » (le public de la sphère publique au sens politique [Lippmann, 2008]). Les sciences sociales de troisième génération ne pourront guère faire autrement que de s'associer à ces plateformes et à ces marques pour produire la science des traces qui seraient alors traitées comme des « réplifications », comme nous le proposons plus loin. Les traces produites étant dépendantes des plateformes, on ne peut guère espérer les modifier à la source. Il est en revanche possible d'exploiter les traces produites par les plateformes en les détournant de l'usage pour lequel elles avaient été conçues (selon le principe de réaffectation ou *repurposing* de Rogers [2013]). La règle veut ici qu'on ne prenne en compte aucune explication d'un autre niveau ou d'un autre monde, mais que l'on puisse comparer des vitesses de propagation, des rythmes ou des transformations éventuelles (par exemple par contamination à d'autres domaines). La différence devra se faire sur la capacité à faire voir des processus qui n'avaient jamais encore été mis en évidence, ou bien en raison des limites techniques prén numériques, ou encore des cibles adoptées par les générations précédentes de sciences sociales.

Par exemple, une étude des fausses informations, comme celle de Vosoughi *et al.*, (2018), permet de montrer que ce n'est pas leur contenu

qui génère leur viralité, mais leur score de nouveauté, car les esprits des humains sont affectés par ce qui choque, ce qui contredit ou qui surprend, et de ce fait, toute nouvelle fantaisiste et incroyable suscite une réaction, quand bien même il s'agit de montrer le caractère invraisemblable de cette nouvelle. Pour cela, les chercheurs ont dû produire un nouvel indicateur, celui de nouveauté. De même, quand on analyse les *tweets* publiés lors des émeutes de Ferguson en 2014 (Srijith *et al.*, 2017), il est possible de retrouver des histoires et des lignes narratives différentes portant ou sur le jeune, ou sur la police et qui, pour certaines, se propagent longtemps et pour d'autres, s'arrêtent assez vite. La comparaison quantifiable dans un contexte similaire « d'éléments de culture » (Dennett, 2017) assez voisins, mais différents permet d'entrer dans les propriétés de ces lignes narratives et fait comprendre ainsi leur succès relatif. Ces exemples se focalisent sur les propriétés des messages, mais d'autres travaux qui font partie intégrante de la théorie des répliques que nous proposons portent sur les propriétés des propagations. Ainsi Cheng *et al.* (2016) ont étudié les formes de propagation d'un mème célèbre (*How to be skinny*) à partir de 12 millions de messages sur Facebook. Ils ont cherché à savoir quelles étaient les conditions de la courbe d'amorçage de la propagation, laquelle permettait de prédire une propagation plus large et plus longue du mème. Ils concluent qu'un mème qui démarre faiblement a peu de chances de connaître des cascades à répétition, tout autant qu'un mème qui connaît un énorme succès dès le début de son lancement, car il semble épuiser toute l'audience potentielle d'un seul coup. En revanche, un succès modéré, mais observable dans des répliques à partir de sources différentes dès le début de la propagation constitue un facteur clé de prédiction d'une récurrence future des cascades. Cela s'explique par le fait que la diversité des milieux pénétrés permet aux différentes versions du mème de s'étendre, ce qui peut prendre du temps et faire apparaître de nouveaux pics d'attention.

Des générations aux modèles conceptuels

Le modèle des trois générations de sciences sociales est un point de départ, et il a été formulé avant tout dans la lignée des travaux d'Alain Desrosières sur la quantification qui constitue les sociétés elles-mêmes. Kitchin et McArdle (2016) proposent une distinction entre données administratives,

données d'enquête et données massives, une classification qui semble ontologique et anhistorique. Elle recoupe cependant entièrement notre proposition de générations de sciences sociales, fondée sur une histoire de la quantification inspirée des études des sciences et technologies (*sciences and technologies studies* ou STS) et qui pointe la différence de point de vue sur le social que chaque génération mobilise. Nous passons ainsi d'une vision technique des données massives et de leurs propriétés listées à l'infini à un tableau structuré des visions possibles du social encapsulées dans les méthodes.

Les différences entre générations ont pu apparaître dans la présentation précédente comme si elles étaient inscrites dans une généalogie alors que ce n'est pas du tout l'approche adoptée, quand bien même un positionnement dans une succession historique a été repéré et un héritage est bien repris et transformé. Toutefois, la coexistence contemporaine des générations doit être rappelée et, de la même façon, l'absence de hiérarchie entre les générations de sciences sociales. En requalifiant les générations de sciences sociales en termes de fréquence (haute, moyenne, basse), il est possible de faire un pas de côté vers une version plus cohérente avec l'expérience même des acteurs : la haute fréquence. Twitter constitue un monde différent par exemple de celui des modes, des opinions et des intentions de vote, qui se cristallisent à intervalles réguliers, de façon cyclique. Cependant, ils sont tous les deux sur une longueur d'onde bien différente encore des transformations démographiques qui chevauchent les générations. Les uns et les autres peuvent devenir numériques ou numérisés, mais cela ne les rend pas pour autant identiques et ce serait une erreur de les regrouper tous dans les sciences sociales computationnelles (« *computational social sciences* ») (Alvarez, 2016), alors que les phénomènes traités sont très différents.

Cependant, au-delà de ces différences de longueur d'onde, c'est une autre approche que nous voudrions développer ici, pour montrer en fait à quel point les différents types de sciences sociales que nous avons mis en évidence recoupent des oppositions classiques entre paradigmes dans les sciences sociales. Nous pensons qu'il est utile de les simplifier en autant de « modèles conceptuels », au sens de Norman (1988). Ce dernier en faisait la condition de l'appropriation des techniques, dans la mesure où tout dépliage complet des dispositifs (ici des méthodes des sciences sociales) avait tendance à agir comme un repoussoir pour les non-spécialistes, qui

se trouvent pris dans les jargons et les éventuelles controverses entre scientifiques. Alors qu'à l'inverse, tout repli complet du dispositif sur une supposée évidence des interfaces intuitives empêchait les usagers et usagères de maîtriser effectivement les outils (l'« effet boîte noire » que les sciences sociales peuvent aussi apprécier parfois et qui aboutit à qualifier abusivement certains énoncés de « science faite »). Le modèle conceptuel constituait pour Norman le bon niveau de présentation des informations pour que des prises (Bessy et Chateauraynaud, 1995) soient possibles. Présenter la diversité des sciences sociales en trois points de vue pourra sembler abusivement réducteur à certaines personnes, et déjà trop perturbant à d'autres. Il nous semble que cela constitue pourtant le niveau des modèles conceptuels que préconisait Norman, qui redonnent prise aux non-spécialistes ainsi qu'aux spécialistes de la science des données (*data scientists*), qui les mobilisent déjà sans le savoir.

Le tableau généalogique des trois générations de sciences sociales (société, opinion et réplifications) peut être présenté en termes de postures qui président à chacune des approches. La tendance à vouloir épuiser tous les comptes rendus du monde par une seule de ces approches est fréquente et génère des guerres de tranchées et des incompréhensions, mais rien ne justifie *a priori* la disqualification de certaines des postures dès lors qu'on peut montrer leur fécondité empirique sur le terrain.

Le tableau des trois générations peut ainsi être prolongé en inscrivant chacune des générations dans une posture ou un point de vue (voir le tableau 2.1).

Les oppositions historiques en sociologie entre structure et préférences individuelles se sont rejouées sans cesse dans le combat entre sociologies d'inspiration marxiste (dont celle de Bourdieu et toutes les sociologies déterministes) et sociologies individualistes (dont celle de Boudon et l'individualisme méthodologique, mais aussi toutes celles qui ont emprunté à la théorie des jeux). Des approches de propagation ont fait leur apparition, mais plus rarement, car les influences de Leibniz ou

TABLEAU 2.1

Inscription des générations dans des postures (1)

Structure	Préférences individuelles	Propagation
Société	Opinion	Réplifications

de Deleuze d'un côté et celles de James et du pragmatisme de l'autre ont été vigoureusement combattues (par Durkheim lui-même), jusqu'à l'apparition de la théorie de l'acteur-réseau de Callon, Latour et Law (Akrich *et al.*, 2006). Le numérique donne une chance d'exister tout à fait inédite à ces approches par la traçabilité qu'il permet, quand bien même la théorie de l'acteur-réseau avait déjà posé les bases d'une méthode proprement sociologique évolutionniste, indépendamment du numérique. L'innovation scientifique et technique était d'ailleurs son terrain privilégié.

Héritages, voisinages et arbitrages

Dans la conclusion du livre *La ville-événement* (Boullier, 2010), traitant de mouvements de foule et de mouvements d'opinion, je proposais de penser les situations de changement social en tenant compte de trois dimensions qui sont autant d'influences ou de rayons imitatifs qui nous traversent et qui constituent le social. Tarde était l'inspirateur de ce modèle, jusque dans sa théorie des croyances et des désirs comme substances imitatives. Je propose donc de distinguer les héritages, les arbitrages et les voisinages.

La théorie de la société de type durkheimien repose ainsi sur une posture qui privilégie la puissance d'agir de la structure, son agentivité (la société comme structure s'impose aux individus) et le poids des héritages y est prépondérant. Bourdieu (1964) systématisera cela en parlant de « capital » et *Les héritiers* n'est pas pour rien le titre de l'ouvrage qui l'a rendu célèbre. Cette reproduction peut être reconstituée en exploitant la continuité de longues séries de données, qui prennent la forme d'ondes longues à basse fréquence.

De même, la question du voisinage est en connexion directe avec l'approche des répliques. Observer des mouvements de foule, c'est mesurer la force de contagion de la proximité physique, comme on le voit dans les mouvements de « ola » dans les stades. Observer le *buzz* à haute fréquence, c'est mesurer la force de contagion de la proximité numérique et de la synchronisation qui (avec les notifications) produit sur le plan temporel l'équivalent de la proximité spatiale. Il s'agit toujours de plusieurs voisinages, et non de celle dont on parle par exemple dans des quartiers urbains. Une modélisation d'effets de voisinage comme celle de Schelling (1960) repose quant à elle sur la simplification extrême des capacités de

décision des entités modélisées : elles quittent leur quartier dès lors qu'un seuil de similarité (noir/blanc) est franchi, sans réellement penser aux propagations entre voisins. Elle relève plutôt de la deuxième génération, c'est-à-dire des préférences individuelles et des capacités de calcul des agents.

Ces deux influences, héritages et voisinages sont en conflit, elles génèrent de l'hésitation, disait Tarde (1895), et dans ce conflit se loge toute la dynamique sociale, qui permet d'observer du changement et non une pure reproduction, mais du changement qui s'installe dans la durée et finit par faire structure et non une agitation permanente, quand bien même la frénésie des *tweets* pourrait le laisser penser.

La question archi-rebattue de « changement et continuité » reviendrait donc en force ici, mais elle n'a guère d'intérêt, si ce n'est de nous permettre d'observer qu'elle provient, d'un point de vue pragmatiste, de véritables luttes d'influence, dont on ne peut comprendre le résultat qu'après coup. En effet, c'est l'hésitation qui doit être aussi formalisée. C'est pourquoi je propose de prendre en compte, au même niveau, les enjeux de l'arbitrage qui s'appuient souvent sur un modèle de l'agent calculateur, idéal de l'économie fétichisée et modélisée avec théorie des jeux à l'appui, mais qui reposent aussi sur des dynamiques de désirs. Effectivement, ce sont les désirs (et donc les peurs) et l'énergie qui est investie dans cet arbitrage qui sont en jeu, comme l'avait dit Tarde (1895). L'arbitrage ne renvoie pas à une prise de décision, mais plutôt à une compétition entre désirs qui peuvent se mesurer et qui exercent alors leur influence sur nos comportements. Le concept de « marché » cherche à saisir ces moments d'arbitrage, ces préférences individuelles, par la mécanique des prix et plus exactement de la monnaie (et non de la valeur) (Orléan, 2011). Cet arbitrage tend à devenir autoréférentiel (pour l'individu, pour la société comme pour la théorie) en particulier dans le capitalisme financier numérique qui est le nôtre, mais les forces des voisinages et des héritages jouent comme des forces de rappel constantes.

TABLEAU 2.2

Inscription des générations dans des postures (2)

Structure	Préférences individuelles	Propagation
Société	Opinion	Réplifications
Héritages	Arbitrages	Voisinages

Dès lors, les postures théoriques ne sont plus incompatibles ni rivales, si l'on admet qu'elles rendent compte de processus d'influences différents, sans espérer rendre compte d'un tout ou d'une combinaison de déterminations. Elles ne sont plus seulement issues des méthodes de quantification et donc d'objectivation mises en œuvre à des époques différentes.

Si le numérique doit permettre un saut, c'est sur le plan du calcul, notamment parce que des degrés de structure, des degrés de préférences individuelles ou de réplifications peuvent être calculés et donc discutés et combinés alors qu'une approche par « essence structurelle » (et non structurale) est obligatoirement binaire (présence ou non). Dans les visions totalisantes, l'agentivité n'est pas vraiment distribuée, elle est entièrement confiée, pour les besoins de l'analyse, à la structure, aux préférences individuelles ou aux réplifications. Le potentiel de classification, d'expérimentation et de validation de ces classes bien connu des informaticiens prend alors tout son sens. Les héritages (et donc les structures) se calculent en degrés d'irréversibilité, ce qui permet de comprendre les effondrements rapides si l'on choisit les bons indicateurs. Le marché et l'arbitrage deviennent des degrés de désirabilité et ni l'utilité ni le travail ne constituent leur référence externe, mais le jeu des attentes réciproques, dont la rivalité mimétique de Girard (1979), pouvait donner une idée (une métrique comme celle de la volatilité d'un titre en bourse peut la rendre calculable). Les réplifications relèvent, elles, de degrés d'immunité (Sloterdijk, 2005) puisque c'est dans la perméabilité des frontières que le voisinage vient perturber les effets des héritages. Selon la porosité plus ou moins grande de ces frontières (notre réceptivité), la circulation des entités (idées, slogans, mêmes) qui nous voient et nous traversent sera plus ou moins facilitée.

La haute fréquence des réseaux numériques demande une théorie des réplifications

L'absence de méthodes empiriques fiables pour certaines approches permet seule d'expliquer l'ignorance ou la prétention hégémonique de chacune des approches. Le seul porteur de l'approche des traces et des propagations, Gabriel Tarde, pourtant quantitativiste qualifié, n'a jamais bénéficié d'une reconnaissance universitaire comparable à celle de Durkheim ou de Weber. Ce n'est pourtant pas le numérique qui fait

émerger des entités inédites, il se contente de rendre visibles et calculables ces entités (les répliques) qui faisaient partie de l'expérience ordinaire (comme les conversations et les rumeurs) et qui avaient été conceptualisées par Tarde (1895) dans sa théorie de l'imitation (qui comprend aussi opposition et invention et ne se limite donc pas à une simple duplication durant la propagation). Ce que nous avons appelé une troisième génération de sciences sociales existait donc potentiellement, mais comme l'ont montré les études des sciences et technologies (STS) et la théorie des conventions, tant que les dispositifs de mesure n'étaient pas disponibles ni étendus pour rendre commensurables les entités en question, elles demeuraient dans les limbes des sciences sociales. L'ère numérique et la traçabilité généralisée qu'elle permet (et qui s'étend désormais aussi aux objets les plus ordinaires) fournit une nouvelle métrique centrée sur les propagations et permet aux sciences sociales, si elles s'en emparent, de dégager un nouveau point de vue sur le social, celui des répliques. Certes, toutes les approches, tous les points de vue et tous les modèles conceptuels peuvent prétendre exploiter la numérisation générale du monde. Toutefois, il serait surprenant que celle qui bénéficie le plus de la visibilité nouvelle fournie à « ses » entités (ce qui circule, les idées et les rumeurs qui nous traversent et nous font agir) ne renforce pas une science sociale spécifique, celle des répliques, encore à inventer.

Un tableau synchronique des approches en question permet de restituer leurs différences et de leur donner toute leur place, sans prétendre à une explication totale à partir d'un point de vue unique. Chacune d'elles est un point de vue sur le monde, sachant qu'il reste impossible de saisir tous les points de vue à la fois (avec toutes les limites de l'analogie, la physique quantique procède ainsi). Ce point de vue dépend entièrement des instruments qu'il mobilise et qui cadrent le monde tout en permettant sa saisie en même temps (Latour, 1990). Néanmoins, ces points de vue se focalisent sur des entités particulières qui ont dans chaque cas un statut privilégié, un pouvoir d'agir, une agentivité, d'emblée privilégiée et distribuée inégalement entre les entités. La focalisation sur les « structures », associée à l'exhaustivité des recensements, diffère nettement de la focalisation sur les « préférences individuelles » associée à la représentativité des sondages. Les résultats sont mis à contribution avant tout dans le marketing, dans les études d'opinion publique, mais aussi dans toutes les grandes enquêtes d'attitude utilisées par les économistes, par exemple sur

TABLEAU 2.3

Trois points de vue en sciences sociales

Point de vue	Structure	Préférences individuelles	Réplifications
Concepts clés	Système, reproduction, patrons	Choix rationnel, prise de décision, théorie des jeux	Événement, incertitude, propagation, chaos, crise
Distribution de l'agentivité (du pouvoir d'agir)	Collectifs, agrégats, communautés, réseaux	Individus, acteurs, stratégies, influenceurs	Entités circulantes, ondes, mêmes
Sources des croyances et des désirs	Héritage	Arbitrage	Voisinage
Longueur d'onde	Ondes longues	Cycles	Ondes à haute fréquence
Analyse de réseaux	Structure du réseau	Nœuds	Flux

le sentiment de bonheur ou de confiance en l'avenir, etc. (par exemple, Algan et Cahuc, 2010). Ces deux entités (structure sociale et préférences individuelles agrégées sous le nom de marché) sont elles-mêmes bien différentes des entités circulantes (Le Béchech et Boullier, 2014), qui peuvent être des messages, des biens ou des êtres qui ont leurs propriétés spécifiques analysées parfois en sémiotique, mais traitées avant tout dans la mémétique.

Le travail des sciences sociales pour penser leur propre activité se trouve ainsi éclairé par la contrainte d'explication propre au numérique. En cherchant à savoir ce que les sciences sociales peuvent réellement faire des traces du numérique, de ces traces nativement numériques rencontrées sur le Web et sur les réseaux sociaux, nous avons bien compris que les problèmes soulevés n'étaient pas du même ordre que ceux posés par les données issues d'autres générations de méthodes de quantification, quand bien même elles sont numérisées (recensements et sondages sont aussi numériques, évidemment). D'autres entités s'imposent dès lors que l'on peut les tracer et les mesurer (quand bien même elles étaient là bien avant !). La priorité nous semble donc de contribuer à construire cette théorie des réplifications, des phénomènes de haute fréquence qui résultent

de voisinages et de circulations d'entités qui demandent à être caractérisées. Néanmoins, en plaçant cette nouvelle piste en perspective historique, nous ne disqualifions aucune des générations précédentes, aucun des modèles conceptuels qui ont fondé les sciences sociales (les structures et les préférences). Et chacune d'elles s'emparera des outils de calcul de l'apprentissage automatique à sa façon, sans qu'il soit nécessaire de chercher à les unifier. Analyser les réseaux de façon structurale ou sur le mode égocentré en privilégiant les nœuds et leur agentivité ne doit pas faire oublier ce qui circule et garde son pouvoir d'agir (et que certaines analyses de réseaux ont eu tendance à sous-estimer). Les méthodes numériques seront donc très différentes selon la distribution d'agentivité que l'on fait dès le départ sans espoir de reconstituer un quelconque « tout » (Latour *et al.*, 2012).

Références

- AKRICH, Madeleine, Michel CALLON et Bruno LATOUR, *Sociologie de la traduction : textes fondateurs*, Presses des Mines, 2006.
- ALGAN, Yann et Pierre CAHUC, « Inherited Trust and Growth », *American Economic Review*, vol. 100, n° 5, 2010, p. 2060-2092.
- ALVAREZ, R. Michael (dir.), *Computational Social Sciences. Discovery and Prediction*, Cambridge University Press, 2016.
- ANDREOU, Athanasios *et al.*, « Investigating ad transparency mechanisms in social media: A case study of Facebook's explanations », dans *Network and Distributed Systems Security (NDSS) Symposium 2018*, San Diego, 2018, p. 1-15.
- BESSY, Christian et Francis CHATEAURAYNAUD, *Experts et faussaires. Une sociologie de la perception*, Métailié, 1995.
- BLONDIAUX, Loïc, *La fabrique de l'opinion. Une histoire sociale des sondages*, Éditions du Seuil, 1998.
- BOULLIER, Dominique, « Big data challenges for social sciences: From society and opinion to replications », *International Sociological Association esymposium*, vol. 7, n° 2, 2017.
- BOULLIER, Dominique, *Sociologie du numérique*, Armand Colin, [2016] 2019.
- BOULLIER, Dominique, « Les sciences sociales face aux traces du Big Data. Société, opinion ou vibrations ? », *Revue française de Science politique*, vol. 65, n° 5-6, 2015, p. 805-828.
- BOULLIER, Dominique, *La ville-événement. Foules et publics urbains*, Presses universitaires de France, 2010.
- BOULLIER, Dominique et El Mahdi EL MHAMDI, « Des modèles aux pratiques: le Machine Learning à l'épreuve de la complexité », *Revue d'anthropologie des connaissances*. Commentaire sur le texte source: 14-1, 2020, en ligne. <http://journals.openedition.org/rac/4260>.

- BOULLIER, Dominique et Audrey LOHARD, *Opinion mining et Sentiment analysis: méthodes et outils*, Open Editions Press, 2012.
- BOURDIEU, Pierre et Jean-Claude PASSERON, *Les héritiers: les étudiants et la culture*, Éditions de Minuit, 1964.
- CHENG, Justin *et al.*, «Do cascades recur?», dans *WWW, 16 Proceedings of the 25th International Conference on World Wide Web*, Montréal, 2016, p. 671-681.
- DAWKINS, Richard, *The Selfish Gene*, Oxford University Press, 1976.
- DENNETT, Daniel, *From Bacteria to Bach and Back. The Evolution of Minds*, Penguin Books, 2017.
- DESROSIÈRES, Alain, *Prouver et gouverner. Une analyse politique des statistiques publiques*, La Découverte, 2014.
- DESROSIÈRES, Alain, *La politique des grands nombres. Histoire de la raison statistique*, La Découverte, 1993.
- DOMINGOS, Pedro, *The Master Algorithm: How the Quest for the Ultimate Learning Machine Will Remake Our World*, Basic Books, 2015.
- EYMARD-DUVERNAY, François *et al.*, «L'économie des conventions ou le temps de la réunification dans les sciences sociales», *Problèmes économiques*, n° 2838, 2004, p. 1-8.
- GIRARD, René, *La violence et le sacré*, Éditions Grasset, 1979 [1972].
- KITCHIN, Rob et Gavin MCARDLE, «What makes Big Data, Big Data? Exploring the ontological characteristics of 26 datasets», *Big Data & Society*, vol. 3, n° 1, 2016, p. 1-10.
- KLEINBERG, Jon, «Bursty and hierarchical structure in streams», dans *Proceedings of the Eighth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, Edmonton, 2002, p. 91-101.
- LATOUR, Bruno, *La science en action*, La Découverte, 1990 [1987].
- LATOUR, Bruno *et al.*, «The whole is always smaller than its parts: A digital test of Gabriel Tarde's monads», *British Journal of Sociology*, vol. 63, n° 4, 2012, p. 590-615.
- LE BÉCHEC, Mariannig et Dominique BOULLIER, «Communautés imaginées et signes transposables sur un "Web territorial"», *Études de communication*, n° 42, 2014, p. 113-125.
- LE BÉCHEC, Mariannig, Dominique BOULLIER et Maxime CRÉPEL, *Le livre-échange. Vies du livre et pratiques des lecteurs*, C & F Éditions, 2018.
- LESKOVEC, Jure, Lars BACKSTROM et Jon KLEINBERG, «Meme-tracking and the dynamics of the news cycle», dans *KDD '09 Proceedings of the 15th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, Association for Computing Machinery, 2009, p. 497-506.
- LIPPMANN, Walter, *Le public fantôme*, Démopolis, 2008 [1925].
- MACKENZIE, Adrian, *Machine Learners: Archaeology of a Data Practice*, The MIT Press, 2017.
- MARRES, Noortje, *Digital Sociology*, Polity Press, 2017.

- MARRES, Noortje, « The issues deserve more credit: Pragmatist contributions to the study of public involvement in controversy », *Social Studies of Science*, vol. 37, n° 5, 2007, p. 759-778.
- MARRES, Noortje et Esther WELTEVREDE, « Scraping the social? Issues in live social research », *Journal of Cultural Economy*, vol. 6, n° 3, 2013, p. 313-335.
- NORMAN, Donald A., *The Design of Everyday Things*, Basic Books, 1988.
- ORLÉAN, André, *L'empire de la valeur*, Éditions du Seuil, 2011.
- ROGERS, Richard, *Digital Methods*, The MIT Press, 2013.
- SCHELLING, Thomas C., *The Strategy of Conflict*, Harvard University Press, 1960.
- SHIFMAN, Limor, *Memes in Digital Culture*, The MIT Press, 2014.
- SLOTERDIJK, Peter, *Sphères III - Écumes*, Maren Sell Éditeurs, 2005.
- SRIJITH, P. K. *et al.*, « Sub-story detection in Twitter with hierarchical Dirichlet processes », *Information Processing & Management*, vol. 53, n° 4, 2017, p. 989-1003.
- TARDE, Gabriel, *La logique sociale*, Félix Alcan, 1895.
- THÉVENOT, Laurent, « Les investissements de forme », dans Laurent THÉVENOT (dir.) *Conventions économiques*, Presses universitaires de France et Centre d'études de l'emploi, Paris, 1986, p. 21-71.
- VOSOUGHI, Soroush, Deb ROY et Sinan ARAL, « The spread of true and false news online », *Science*, vol. 359, n° 6380, 2018, p. 1146-1151.

CHAPITRE 3

L'analyse du numérique dans une perspective complexe avec la sociologie des usages

Alexandre Coutant, Jean-Claude Domenget

Comment s'assurer de la cohérence entre les ambitions scientifiques affichées en sociologie des usages et la conduite concrète des terrains ? Comment suivre le principe décrit par Durkheim en 1887, afin de concilier les approches micro et macro : « Il n'y a qu'une manière de parvenir au général, c'est d'observer le particulier : non pas superficiellement et en gros, mais minutieusement et par le détail » ? Inscire une recherche en sociologie des usages revient à faire le choix d'approches et de théories ayant l'ambition de conserver toute leur complexité aux phénomènes étudiés (Denouël et Granjon, 2011 ; Jauréguiberry et Proulx, 2011). Si cette perspective s'avère stimulante intellectuellement, elle reste délicate à mettre en œuvre au sein d'une recherche empirique. En effet, les protocoles, dimensions à prendre en compte et étapes s'en trouvent peu balisés. Nous proposons à cette fin un guide fondé sur notre propre expérience de recherche en sociologie des usages. Au cours de ses quarante années d'existence, cette sociologie a multiplié les propositions de programmes ambitieux et les invitations à prêter attention à des dimensions trop peu visibles, bien que déterminantes, de nos rapports aux objets techniques. Sans prétention à l'exhaustivité¹, évoquons la double médiation (Jouët,

1. Pour une analyse plus systématique des différents apports façonnant l'approche de la sociologie des usages, voir Coutant et Domenget (2014).

1993), la sociopolitique des usages (Vitalis, 1994), l'intégration sociale des TIC (Mallein et Toussaint, 1994), la sociologie critique des usages (Granjon, 2004) ou encore la construction sociale des usages (Proulx, 2005). Sans entrer dans le détail de ces propositions, insistons sur ce qui les réunit : la volonté de saisir les multiples dimensions du phénomène, à la fois dans la complexité de son actualisation contemporaine et dans son épaisseur historique, en insistant particulièrement sur le sens pris par ces phénomènes pour les individus.

L'ambition est séduisante, mais les chercheuses et chercheurs peuvent se trouver démunis face à la diversité des terrains que cette appréhension supposerait d'accumuler. L'ampleur de la littérature scientifique associée à ces différentes dimensions peut aussi paraître dépasser largement les capacités de lecture personnelles. En outre, faire dialoguer les apports respectifs de ces littératures implique une érudition et une connaissance interdisciplinaires que l'on imagine mal maîtriser dès ses premiers pas dans la carrière scientifique. Plusieurs critiques ne se sont pas trompées en pointant l'écart entre les ambitions affichées et les études de terrain, se cantonnant souvent à des entretiens, voire à l'analyse d'échanges sur un seul dispositif numérique (Jauréguiberry, 2008; Jouët, 2011). Comment alors s'inscrire dans cette démarche ambitieuse dans les limites d'un travail de mémoire ou de thèse ? Il s'agit de situer sa contribution au sein des différentes étapes d'une enquête d'usages. Convaincus qu'une compréhension des phénomènes d'appropriation des objets techniques communicationnels relève plus d'un programme collectif d'une communauté de scientifiques que de travaux isolés, nous espérons ainsi aider à la cumulativité des travaux menés au sein de la sociologie des usages.

Nous proposons des pistes concrètes pour mettre en cohérence toutes les dimensions d'une recherche sur les usages dans une perspective complexe. Nous illustrons cette présentation par un cas de recherche menée sur les refus de prescription algorithmique. Cette formalisation nous a été nécessaire pour mieux nous représenter en quoi nos bricolages (Becker, 2002) permettaient de répondre aux exigences scientifiques. Il nous semble effectivement que la formation à la recherche insiste peu sur le lien entre les problématisations ou lectures et les terrains ainsi que sur les manières de faire dialoguer les retours de terrains avec les connaissances accumulées au contact de la littérature. Les cours méthodologiques enseignent comment bien appliquer une méthode d'enquête particulière, mais rare-

ment pourquoi ou en quoi ces choix se doivent d'être en cohérence avec la problématisation et le cadre théorique. Les cours théoriques rendent compte des apports d'auteurs reconnus, mais en disent souvent peu sur les modalités de leurs applications à nos propres terrains ou comment ces derniers permettent de les discuter. L'alternance entre les phases d'observation, d'interprétation et de théorisation demeure peu explicitée alors qu'elle constitue le cœur de la richesse des approches qualitatives (Anadon et Savoie-Zajc, 2009). Si ce cadre permet de mieux connecter ces domaines que rien ne devrait isoler, alors il aura rempli ses objectifs. Dans un premier temps, nous présenterons le contexte du projet de recherche qui nous servira d'illustration. Nous détaillerons ensuite les différentes phases du guide proposé en prenant soin de préciser comment elles se sont actualisées dans notre cas.

Le refus d'algorithmes complexes de classement de l'information

Le programme de recherche qui nous sert d'illustration concerne l'acceptation ou non par les internautes de la mise en place, sur les plateformes numériques, d'organisations algorithmiques élaborées de l'information. Toute organisation de l'information passe par un algorithme dans le sens de « série d'instructions permettant d'obtenir un résultat » (Cardon, 2015, p. 7). Certaines organisations nous sont facilement compréhensibles : c'est le cas des classements antéchronologiques organisant les informations de la plus récente à la plus ancienne. Ainsi, les adeptes d'Instagram ont longtemps utilisé la plateforme en remontant leur fil d'informations jusqu'au dernier contenu consulté lors de leur session précédente. Ils s'assuraient ainsi d'avoir bien passé en revue l'intégralité des nouveaux contenus. Par organisations élaborées, nous entendons des logiques de classement et de mise en visibilité qui dépassent la compréhension intuitive, par exemple la mise en place d'algorithmes se fondant sur l'exploitation d'un ensemble opaque de traces de nos navigations et positions antérieures, comme le propose Google depuis longtemps avec sa personnalisation des résultats.

La prolifération des systèmes de recommandation reposant sur des algorithmes élaborés constitue un exemple particulièrement intéressant de phénomènes complexes. Netflix, Google, Amazon ainsi que presque tous les médias socionumériques en implantent peu à peu. Ils sont accom-

pagnés d’imaginaires multiples, tantôt idéalistes, tantôt apocalyptiques (Ertzscheid, 2017). Ces algorithmes se voient même prêter des caractéristiques empruntant aux fonctions les plus élaborées de la cognition humaine (intelligence, autoapprentissage, etc., voir notamment les travaux de Abiteboul et Dowek, 2017). Leur ensemble relativement réduit de logiques de fonctionnement est appliqué à une variété importante de domaines de pratiques (consommation, travail, sociabilité) (Cardon, 2015). Cette tendance amène à placer les individus dans un écosystème informationnel très opaque, qu’ils comprennent peu et dont ils doutent qu’il soit développé prioritairement dans leur intérêt. En découlent de nombreuses controverses dont les ressorts éclairent des enjeux fondamentaux interrogés par la sociologie des usages : formes et portée des prescriptions, confiance, rapport entre types d’expertises – certifiée ou profane, littératies informationnelles et techniques, divergence entre l’usage configuré dans le dispositif et l’usage effectif, confrontation entre routines mises en place par les usagers, usagères et logiques encouragées par les algorithmes, enjeux de qualification et d’évaluation des contenus partagés. C’est la compréhension de ce phénomène qui anime notre projet de recherche. Nous le détaillerons pour illustrer les phases de notre guide et voir comment elles nous amènent à prêter attention à un ensemble varié de plateformes et d’usages.

Une recherche complexe menée de façon concrète

« Tout ce qui est simple est faux, mais tout ce qui ne l’est pas est inutilisable » (Valéry, 1941). Focaliser son analyse sur un ensemble réduit de dimensions est un moyen de simplifier son travail, mais cette simplification se fait au prix de la prise en compte du phénomène dans toute sa densité (Latzko-Toth *et al.*, le présent ouvrage). Le guide que nous présentons cherche à résoudre ce problème. Il propose une mise en visibilité à la fois des dimensions à aborder et de leur enchaînement. Trois phases de la recherche sont distinguées (la problématisation, le terrain et la théorisation), chacune étant composée de trois strates de recherche (les objectifs, les dimensions observées, les exigences à respecter). Le guide peut être représenté synthétiquement par le tableau suivant :

FIGURE 3.1

Guide d'analyse des usages des dispositifs sociotechniques d'information et de communication

	Problématisation	Description	Théorisation
Objectifs	<ul style="list-style-type: none"> • Déconstruction critique des définitions courantes de l'objet • Redéfinition de l'objet 	<ul style="list-style-type: none"> • Appréhension fine de l'objet • Visée de saturation des différentes formes prises par les phénomènes sociotechniques inclus dans l'objet 	<ul style="list-style-type: none"> • Synthèse à visée de généralisation fondée sur la description de l'objet • Explication éclairée des enjeux individuels, sociaux, économiques et politiques de l'objet
Dimensions observées	<ul style="list-style-type: none"> • Revue de littérature • Veille • Analyse du discours d'accompagnement 	<ul style="list-style-type: none"> • Observations de l'activité des usagers • Recueil de leurs représentations de leur activité • Identification des contextes d'usage • Analyse des affordances des dispositifs 	<ul style="list-style-type: none"> • Mise en commun des analyses • Comparaison avec travaux complémentaires • Comparaisons interdisciplinaires
Exigences	<p><i>Accumulation théorique</i></p> <ul style="list-style-type: none"> • Étudier les concepts convoqués • Replacer les problématiques dans le long terme de l'étude des dispositifs sociotechniques 	<p><i>Description</i></p> <ul style="list-style-type: none"> • Multiplier des angles d'éclairage de l'objet • Employer des méthodologiques complémentaires • Déterminer les constantes et les nouveautés dans les usages observés 	<p><i>Ouverture théorique et disciplinaire</i></p> <ul style="list-style-type: none"> • Éprouver les résultats d'enquête auprès d'autres travaux • Relier les usages observés à des pratiques sociales plus larges • Construire une appréhension complexe de l'objet sur la base des différents angles d'éclairage mis à jour
Phases du programme de recherche			

Phase 1 : La problématisation et la définition scientifique d'un objet d'étude

La première étape, consacrée à la problématisation, a pour ambition de s'assurer de la construction d'un objet de recherche détaché des imaginaires sociaux accompagnant une technologie en émergence. Elle a aussi comme objectif de prendre de la distance à l'égard des attentes commerciales ou politiques pouvant s'associer aux incitations à effectuer des recherches autour de cette technologie (Robert, 2016). Le principe consiste à sortir des discours d'accompagnement, non pas pour dire ce que les technologies devraient être *a priori*, dans une optique normative, mais pour problématiser les enjeux associés à ces dernières selon une démarche scientifique.

Cet objectif peut nécessiter un travail important. Il implique un effort de conceptualisation des termes utilisés en fonction des notions abordées et des travaux antérieurs mobilisables. La finalité est d'éliminer les notions-écrans² pour s'assurer que les concepts employés correspondent à des définitions précises qui aideront à circonscrire et à analyser l'objet. Le risque des notions-écrans consiste effectivement à revêtir un caractère d'évidence qui dissuade de s'interroger sur ce qu'elles veulent dire précisément. La conceptualisation conduit à proposer des définitions rigoureuses et exploitables analytiquement de termes employés par les actrices et acteurs impliqués dans le phénomène (ex. : réseaux sociaux, émotion, engagement, participation, intelligence artificielle, données massives, etc.). Elle peut aussi amener à préférer un néologisme ou l'emploi d'un attirail conceptuel préexistant (ex. : système technique, dispositif, plateforme, industrie culturelle, aliénation, activité, transduction, confiance, pratique sociale, etc.). Concrètement, le travail repose sur une revue de la littérature scientifique, une veille sur les termes convoqués à propos de l'objet et sur une analyse des discours d'accompagnement. Quelques enquêtes exploratoires peuvent aussi servir à mettre en lumière d'éventuelles contradictions ou apories dans la manière de traiter le phénomène.

Cette première étape suppose une certaine accumulation théorique. Elle se caractérise par un effort de renouvellement des cadres conceptuels, pour déterminer, par exemple, si le terme « intelligence » permet bien de

2. Miège (2012) qualifie ainsi l'emploi incontrôlé de termes issus des discours politiques, médiatiques ou commerciaux, accompagnant un phénomène et face auxquels il préconise un travail critique de conceptualisation.

penser l'intelligence artificielle ou si une nouvelle conceptualisation est nécessaire (il sera alors utile de s'approprier au préalable la littérature produite par les spécialistes du domaine en psychologie et en sciences cognitives). Les exigences de l'étape de problématisation vont conduire à inscrire l'objet de recherche dans le long terme des études sur les techniques et les supports de communication. La sociologie des usages nous invite en effet à analyser les filiations d'usage entre dispositifs. Il s'agit ainsi de prendre en compte le phénomène analysé au sein d'une pratique sociale dépassant le simple usage *hic et nunc* d'un dispositif, et de prêter attention aux différents supports sur lesquels repose la pratique afin de comprendre ce qu'un nouveau support, avec ses spécificités sociotechniques, change ou reproduit d'une configuration sociotechnique existante. Une telle problématisation donnera la possibilité d'une véritable analyse sociotechnique dans la phase suivante et servira de base à une théorisation. Ajoutons que, dans la perspective d'une démarche inductive valorisée en sociologie des usages, des retours vers cette première étape peuvent être effectués, en fonction des retours des terrains.

Dans notre cas, l'appellation « moteurs de recommandation » et les notions associées (personnalisation, algorithme, données massives, etc.) jouissaient déjà en 2013, lors du démarrage de notre enquête, d'un discours d'accompagnement vantant la pertinence de ces systèmes, leur valeur économique, mais aussi leur grande complexité. Les éléments constituant un impensé paraissaient donc bien établis. Il s'agissait alors, dans cette phase, d'aboutir à une définition plus précise de l'objet en question à l'aide de concepts offrant des angles permettant de l'analyser. Il s'agissait aussi, au-delà de la variété des cas de moteurs de recommandation, de relever des éléments récurrents permettant de ne pas se cantonner à des études de cas isolées, mais bien d'interroger le phénomène comme illustratif d'un fait social plus large (Lahire, 2006), ici, le passage aux algorithmes de la tâche d'organisation de l'information.

Grâce à notre revue de la littérature des travaux reliés à la recommandation, nous avons pu dégager une approche très focalisée sur la qualité de l'information prise en elle-même, souvent endossée par les concepteurs (Drumond, 2016; Kembellec *et al.*, 2014). Cette revue a aussi soulevé la pertinence d'approches plus orientées vers les publics (Esquenazi, 2009; Morelli *et al.*, 2016) et permis de trouver dans le concept de prescription (Gaudin et Hatchuel, 2002; Stenger, 2011) une définition précise des

dimensions engagées dans les contextes de recommandation. Ces éléments conceptuels ont guidé l'élaboration d'une problématique davantage orientée vers les enjeux communicationnels que vers la seule attention aux contenus et à leurs modes de classement : comment les profanes accueillent-ils ces systèmes, se les approprient-ils et forment-ils des représentations à leur égard ?

Phase 2 : La description et l'observation fine du terrain

La deuxième étape consiste à décrire finement le phénomène en question. Soulignons que c'est sur cette phase que se concentre la plus grande partie des critiques adressées à la sociologie des usages, c'est-à-dire que ce courant se cantonnerait à cette description minutieuse, sans avoir au préalable mis en place une problématisation permettant de voir au-delà de « l'économie interne des usages » (Jauréguiberry, 2008) et sans préoccupation suffisante pour le fait de relier ces descriptions à des phénomènes macrosociaux (Jouët, 2011). L'apport souhaité de notre guide consiste à mieux déterminer le rôle de cette phase et à permettre de la rattacher aux autres, en mettant en lumière la façon dont elle les nourrit. Davantage qu'une approche purement inductive, nos revues de littérature laissent plutôt percevoir en sociologie des usages ce qu'Anadon et Savoie-Zajc (2009) ou Catellin (2004) qualifient d'abduction, c'est-à-dire une démarche de fertilisation croisée entre théorisation à partir des matériaux accumulés et déductions permises par le savoir accumulé. La description est donc centrale pour affiner la problématisation et permettre la théorisation. Cette clarification de l'apport de la phase descriptive nous apparaît d'autant plus essentielle que la plupart des travaux de recherche qui découlent de la sociologie des usages se situent généralement dans cette phase.

L'objectif principal de cette deuxième étape est de combiner les résultats de terrains complémentaires dans le but de déceler les constantes et les nouveautés dans les pratiques analysées. Ce cumul doit permettre de relier le cas étudié à des filières d'usages et d'insérer ces usages dans le contexte plus large de pratiques sociales (Jouët, 1993). Entre usages prescrits, usages imaginés et usages réels, les distinctions sont importantes et doivent être étudiées avec précision. Les dimensions complémentaires d'analyse à prendre en compte sont nombreuses : représentations qu'ont les usagers et usagères de leurs activités afin d'approcher les significations

d'usage; observation des pratiques qui intègrent des habitudes fortement ancrées, des discours d'accompagnement, des normes d'(il)légitimité de certaines pratiques; détermination des contextes d'usage qui établissent le sens social que ces derniers prennent; incitations contenues dans les scripts des objets techniques; analyse des stratégies économiques ou politiques des acteurs impliqués, etc. Le modèle méthodologique proposé par De Certeau (1980) fournit une synthèse précieuse de ces dimensions à faire dialoguer. L'analyse de l'activité des individus nécessite que l'on accumule du matériau concernant le donné (c.-à-d. les éléments mis en place par les stratèges tels que l'offre disponible, les interfaces et terminaux, l'organisation de l'information, les fonctionnalités, etc.), les actions (c.-à-d. les performances effectives avec ce donné: le contexte d'usage, les interactions avec les différents actants en ligne et hors-ligne, les fonctionnalités employées et délaissées) et les significations d'usage (c.-à-d. le sens que prennent ces performances pour les individus: finalité de l'activité, conscience des actants impliqués, représentations de la technique et des actants impliqués, normes et valeurs convoquées, représentation du collectif et de ses raisons d'existence, autres significatifs impliqués, etc.).

La variété des problématisations entraîne une forte diversité des points sur lesquels se pencher, ainsi que des modes de collecte privilégiés: formes d'observations, suivis de pratiques sur le long terme, analyse de discours, sémiotique des interfaces, suivis de logs, monographie socio-économique de plateforme, reconstitution de filière technique, carnets de pratiques, entretiens individuels, de groupe ou d'explication, exercices autoréflexifs individuels ou collectifs, jeux de rôle, récits de vie, méthode des portraits, collages et photomontages, évaluation de scripts, analyse historique dans un secteur d'activité, etc. L'exigence principale de cette phase d'observation consiste néanmoins dans la saturation des angles de questionnement soulevés en problématisation. Chaque matériau peut ainsi être justifié par sa contribution pour aborder un ou des éléments dégagés dans cette première phase.

Notre interrogation sur les moteurs de recommandation s'étant précisée autour des cas de refus ou d'indifférence à leur égard, nous avons commencé par recenser les controverses en ligne ayant accompagné leur mise en œuvre. Ces controverses s'avèrent très riches en matériaux complémentaires: prises de paroles argumentées sur différentes plateformes (sur les services eux-mêmes ou par l'intermédiaire de blogues ou vlogues),

tutoriaux pour désactiver ou obfusquer³ l'algorithme, captures d'écran d'exemples d'inefficacité des résultats, formes virales d'humour décrédibilisant le système, discours de promotion puis de défense de l'algorithme par la plateforme. Cependant, en centrant notre attention sur nos façons bien plus ordinaires de négocier avec les dispositifs sociotechniques permises par notre problématisation, nous avons été amenés à compléter ce matériau. Il ne rend effectivement compte que de la faible minorité des personnes prenant la parole en ligne. Pour mieux comprendre les failles du contrat de communication censé lier les plateformes à la variété de leurs publics, nous avons donc rassemblé de la littérature sur les formes d'informations recherchées et d'interactants les appuyant selon les plateformes. Nous avons aussi effectué une analyse sémiotique des moteurs de recommandation de deux plateformes sur lesquelles nous accumulions des terrains et qui se heurtaient à des refus de prescription : Facebook et Twitter. Cette analyse a été orientée par les dimensions identifiées lors de la prescription. Elle a permis de préciser le type d'actions ou de contenus suggérés (l'énoncé dans une perspective discursive), mais aussi le contenu discursif justifiant la pertinence de ces contenus et actions (l'énonciation). Ceci a permis de mettre à jour des inadaptations dues, chez Facebook, à l'inadéquation entre ces contenus ou actions prescrites et les activités effectivement menées et, chez Twitter, à une mauvaise contextualisation de la recommandation, qui empêchait les personnes en faisant usage de comprendre son intérêt.

Phase 3 : La théorisation

La troisième étape relève d'un objectif de théorisation. Celle-ci peut prendre deux formes principales : la première consiste à effectuer une montée contrôlée en généralité. Par exemple, la vaste littérature concernant la participation en ligne a permis de mener à des travaux de synthèse tentant de mettre en évidence ses caractéristiques récurrentes et son inscription dans la pratique sociale plus large de la participation politique (Denouël *et al.*, 2014). Cette première forme de théorisation consiste à effectuer le lien entre les études de cas micro et ce qu'elles disent d'un

3. On parle d'obfuscation pour évoquer toutes les tactiques des individus visant à se rendre moins transparents et à altérer le fonctionnement de la surveillance par les plateformes.

phénomène macro plus large lorsqu'elles sont prises comme un corpus complémentaire d'enquêtes (Proulx, 2005). Précisons si nécessaire que nous ne passons pas forcément par cette étape lorsque nous effectuons une enquête : il peut être parfois prématuré de quitter le niveau descriptif. La seconde forme, plus fréquente, consiste à discuter des théories existantes ou à proposer de nouveaux concepts les complétant. C'est le cas de la proposition qui suggère d'employer le qualificatif « socionumérique » pour décrire les plateformes émergeant avec le développement des techniques du Web 2.0 afin d'éviter le double écueil de ne retenir que la dimension de réseau social ou de réseau numérique (Stenger et Coutant, 2011). La discussion/conceptualisation s'appuie sur les terrains et lectures accumulés, dans une perspective empruntant à nouveau à la fois à l'induction (émergeant des terrains) et à la déduction (découlant de cadres conceptuels existant) (Anadon et Savoie-Zajc, 2009; Catellin, 2004). Précisons que chaque enquête ne débouche pas nécessairement sur une discussion ou une nouvelle conceptualisation. Elle peut illustrer dans un nouveau contexte des éléments déjà connus. Il ne sert à rien de vouloir forcer la discussion théorique; une description dans un nouveau contexte a déjà valeur de confirmation d'une théorie existante et sert ainsi sa montée en généralité.

La théorisation repose sur la mise en commun des différentes dimensions de l'objet observé. Elle doit permettre de fournir un éclairage complexe des enjeux individuels, sociaux, économiques, politiques et culturels liés aux formes contemporaines des dispositifs sociotechniques étudiés. C'est ici qu'interviennent les comparaisons avec des travaux complémentaires ou avec d'autres disciplines qui s'intéressent aux mêmes objets. Pour sortir de ce qui serait un « médiacentrisme », cette théorisation doit rapprocher les usages déterminés localement des autres contextes sociaux, numériques ou non, dans lesquels se joue une pratique sociale plus large, dont l'usage ne constitue qu'une actualisation située, par exemple le maintien de liens d'amitié, le partage de contenus autour de passions, etc. La tripartition données/actions/significations évoquée *supra* avec la figure tutélaire qu'est De Certeau aide de nouveau à prendre en considération les dimensions qui doivent être mises en dialogue pour avancer dans cette phase.

La portée éventuelle des résultats de cette phase peut être évaluée à trois niveaux : une intelligibilité améliorée d'une situation sociale, une

discussion interdisciplinaire renouvelée de cadres théoriques historiques ou la proposition de nouvelles conceptualisations qui permettront de problématiser autrement de nouveaux terrains.

Notre projet de recherche est encore en cours. Nous ne présentons ici que les pistes de théorisation nous paraissant les plus intéressantes à creuser, tout en précisant que celles-ci ne sont pas nécessairement totalement abouties. Le résultat le plus tangible de ce projet se situe sur le plan épistémologique, puisque c'est grâce à ce dernier que nous avons expérimenté le guide que nous venons de présenter. Il consistait dans la mise en œuvre concrète de réflexions menées au sein de plusieurs séminaires rattachés à la généalogie de la sociologie des usages et aux liens avec les autres courants d'analyse des phénomènes sociotechniques (Coutant et Domenget, 2014). Un deuxième résultat relevant de cette phase consiste dans la démonstration de la valeur d'une analyse des usages tirés de l'ordinaire pour se départir des imaginaires associés aux technologies. Au-delà des revendications d'efficacité des systèmes de recommandation et des focalisations sur des résistances spectaculaires, c'est une multitude silencieuse d'activités indifférentes à ces derniers, ou se bricolant des chemins alternatifs, qui ont été rendues visibles. Cet écart non militant et revendicatif constitue certainement le rapport le plus fréquent aux tentatives de prescriptions nous entourant en ligne et cette hypothèse a pu être testée dans d'autres contextes prescriptifs (Trudelle, 2018). Enfin, la revue de littérature nous a permis de constater la précision du cadre conceptuel proposé par Gaudin et Hatchuel (2002) pour aborder la prescription. Cependant, orienté vers la description de ses formes, ce cadre insiste moins sur la circulation de celles-ci au sein de l'espace social. Des efforts sont actuellement déployés pour nous amener à tenter d'associer ce cadre conceptuel aux théories des usages et des publics, notamment au sein du concept de contrat de communication tel qu'il a été pensé par Verón (1988).

Une orientation et non l'imposition d'une manière de faire

Le guide présenté dans ce chapitre vise à aider à situer un travail de recherche au sein des programmes ambitieux mis en avant par la sociologie des usages. Le dévoilement des phases de problématisation, de terrain et de théorisation permet tout d'abord de conserver un fil directeur au sein de la grande variété des cadres conceptuels, méthodologies, rapports

aux terrains et manières de problématiser les phénomènes sociotechniques dans la littérature en sociologie des usages. L'impression de vertige que cette production pléthorique peut donner est ainsi combattue : la connaissance de la totalité de la production importe moins que la cohérence que l'on donne à ces phases et à leur enchaînement. Elles permettent aussi de concevoir nos propres enquêtes comme des contributions à l'une ou l'autre de ces phases, selon l'ampleur des projets. Les critiques d'empirisme acritique ou de théorisation prématurée peuvent ainsi être évitées grâce à la précision de la portée des descriptions et de la production collective sur laquelle repose la montée en généralité ou conceptualisation.

Rappelons cependant qu'il n'y a qu'un pas à franchir pour que le guide se transforme en contrainte. La présentation par étapes successives peut en effet donner la fausse impression d'une progression linéaire que les projets de recherche ne suivent pas systématiquement, loin de là. Il convient de ne pas prendre ces phases pour des étapes hermétiquement séparées. Si le guide est utile pour envisager le déroulement général de la recherche, au quotidien, l'on peut être amené à revenir ou à avancer dans certaines phases, selon ce que nos lectures ou terrains nous suggèrent. Le respect de la démarche qualitative exigée par la sociologie des usages, avec tout ce qu'elle exige de mobilité entre les périodes de problématisation, de description et de théorisation, demeure effectivement la priorité. Loin d'être figées, ces étapes peuvent donc se chevaucher chronologiquement ou nécessiter des allers-retours entre littérature scientifique et terrain.

Ce guide possède la double vertu de spécifier la portée des résultats, mais aussi d'éviter de jauger ceux-ci selon des exigences relevant d'une autre étape que celle dans laquelle ils s'inscrivent. Il aide à projeter les projets de recherche dans des temporalités longues et à organiser la mise en cohérence des terrains. Il est aussi un outil précieux pour expliquer l'apport de chacune des publications. Cette clarification nous paraît très utile pour relever collectivement les ambitions des programmes en sociologie des usages. Elle permet d'éviter les incompréhensions au sein des sciences de l'information et de la communication entre des traditions de recherche différentes qui s'intéressent aux liens entre technique et société.

Références

- ABITEBOUL, Serge et Gilles DOWEK, *Le temps des algorithmes*, Éditions le Pommier, 2017.
- ANADON, Marta et Lorraine SAVOIE-ZAJC, « Introduction: L'analyse qualitative des données », *Recherches qualitatives*, vol. 28, n° 1, 2009, p. 1-7.
- BECKER, Howard S., *Les ficelles du métier: comment conduire sa recherche en sciences sociales*, La Découverte, 2002.
- CARDON, Dominique, *À quoi rêvent les algorithmes. Nos vies à l'heure des big data*, Éditions du Seuil, 2015.
- CATELLIN, Sylvie, « L'abduction: une pratique de la découverte scientifique et littéraire », *Hermès, La Revue*, vol. 2, n° 39, 2004, p. 179-185.
- COUTANT, Alexandre et Jean-Claude DOMENGET, « Un cadre épistémologique pour enquêter sur les dispositifs sociotechniques d'information et de communication », dans Hélène BOURDELOIE et David DOUYÈRE (dir.), *Méthodes de recherche sur l'information et la communication*, Mare et Martin, 2014.
- DE CERTEAU, Michel, *L'invention du quotidien, tome 1: Arts de faire*, Union générale d'édition, 1980.
- DENOUEËL, Julie et Fabien GRANJON (dir.), *Communiquer à l'ère numérique: regards croisés sur la sociologie des usages*, Presses des Mines, 2011.
- DRUMOND, Gabrielle Silva Mota, *La configuration des usages sur Netflix: le système de recommandation Cinematch et la représentation de l'utilisateur*, mémoire de maîtrise en communication, Université du Québec à Montréal, 2016.
- DURKHEIM, Emile, « La science positive de la morale en Allemagne », *La Revue philosophique*, n° 24, 1887, p. 33-142 et 275-284.
- ERTZSCHEID, Olivier, *L'appétit des géants: pouvoir des algorithmes, ambitions des plateformes*, C & F Éditions, 2017.
- ESQUENAZI, Jean-Pierre, *Sociologie des publics*, La Découverte, 2009.
- GAUDIN, Thierry et Armand HATCHUEL, *Les nouvelles raisons du savoir: vers une prospective de la connaissance*, Éditions de l'Aube, 2002.
- GRANJON, Fabien, « De quelques éléments programmatiques pour une sociologie critique des usages sociaux des TIC », *intervention au sein de la journée d'étude organisée par le LARES-Université de Rennes 2, sous la direction de Smail Hadj-Ali, « Les rapports société-technologie du point de vue des sciences de l'homme et de la société »*, 2004.
- JAURÉGUIBERRY, Francis, « Sociologie des usages des technologies de la communication. "L'école française" des années 80 », dans *Actes des travaux du Comité de recherche « Sociologie de la communication »*, XVIII^e Congrès international des sociologues de langue française, 2008, p. 12-18.
- JAURÉGUIBERRY, Francis et Serge PROULX, *Usages et enjeux des technologies de communication*, Érès, 2011.
- JOUËT, Josiane, « Des usages de la télématique aux Internet studies », dans Julie DENOUEËL et Fabien GRANJON (dir.), *Communiquer à l'ère numérique: regards croisés sur la sociologie des usages*, Presses des Mines, 2011, p. 4590.

- JOUËT, Josiane, « Pratiques de communication et figures de la médiation », *Réseaux*, vol. 11, n° 60, p. 99-120, 1993.
- KEMBELLEC, Gérard, Ghislaine CHARTRON et Imad SALEH (dir.), *Les moteurs et systèmes de recommandation*, ISTE Éditions, 2014.
- LAHIRE, Bernard, « Nécessité théorique et obligations empiriques », *Revue du MAUSS*, vol. 1, n° 27, 2006, p. 444-452.
- MALLEIN, Philippe et Yves TOUSSAINT, « L'intégration sociale des TIC : une sociologie des usages », *Technologies de l'information et société*, vol. 6, n° 4, 1994, p. 315-335.
- MIÈGE, Bernard, *L'édification des SIC : encore et toujours*, conférence introductive au congrès de Rennes de la SFSIC, 2012.
- MORELLI, Pierre, Nathalie PIGNARD-CHEYNEL et Didier BALTAZART, *Publics et TIC : confrontations conceptuelles et recherches empiriques*, Éditions universitaires de Lorraine, 2016.
- PROULX, Serge, « Penser les usages des technologies de l'information et de la communication aujourd'hui : enjeux, modèles, tendances », dans Lise VIEIRA et Nathalie PINÈDE (dir.) *Enjeux et usages des TIC : aspects sociaux et culturels*, tome 1, Presses universitaires de Bordeaux, 2005, p. 720.
- ROBERT, Pascal, *L'impensé numérique, tome 1 : Des années 1980 aux réseaux sociaux*, Éditions des Archives contemporaines, 2016.
- STENGER, Thomas, « La prescription de l'action collective », *Hermès, La Revue*, vol. 1, n° 59, 2011, p. 127-133.
- STENGER, Thomas et Alexandre COUTANT, « Introduction », *Hermès, La Revue*, vol. 1, n° 59, 2011, p. 917.
- TRUELLE, Camille, *Sous quelles conditions les abonnés d'un instagrameur intervenant sur l'alimentation acceptent que celui-ci s'associe à des marques?*, mémoire de maîtrise en communication, Université du Québec à Montréal, 2018.
- VALÉRY, Paul, *Mauvaises pensées et autres*, Les Cahiers du Sud, 1941.
- VERÓN, Eliseo, *La sémiologie sociale : fragments d'une théorie de la discursivité*, Presses universitaires de Vincennes, 1988.
- VITALIS, André (dir.), *Médias et nouvelles technologies : pour une socio-politique des usages*, Éditions Apogée, 1994.

DEUXIÈME PARTIE

Les méthodes ethnographiques

CHAPITRE 4

L'ethnographie des communautés en ligne et des médias sociaux : modalités, diversité, potentialités

*Christine Hine*¹

La reconnaissance de la richesse et de la diversité des interactions sociales rendues possibles par Internet est indissociable de l'avènement des méthodes ethnographiques qui documentent ces interactions et examinent leurs répercussions. La tradition ethnographique établie, qui considère l'ethnographe comme un instrument de recherche incarné, s'est transportée aux espaces sociaux du Web. L'importance que revêt en ethnographie l'étude détaillée et approfondie des pratiques sociales et du sens que leur prêtent les individus a permis aux ethnographes de décrire ce que font réellement les internautes dans des situations précises. Les enquêtes ethnographiques menées en ligne ont contribué à établir la nature éminemment sociale des interactions qui se déploient dans les mondes numériques et à étudier les nouvelles formations culturelles qui en émergent. Plusieurs approches sont donc apparues dans le champ de l'ethnographie en ligne à mesure qu'Internet s'est développé. Elles soulèvent des enjeux méthodologiques spécifiques qui doivent être minutieusement examinés au moment d'appliquer les principes de la recherche ethnographique à des terrains en ligne.

1. Texte traduit de l'anglais par Margot Lacroix, avec la collaboration de David Myles et Guillaume Latzko-Toth, d'abord publié sous le titre de « Ethnographies of online communities and social media: modes, varieties, affordances », dans *The SAGE Handbook of Online Research Methods*, SAGE Publications, 2017.

Au-delà des décisions à prendre quant à la démarche à mettre en œuvre sur un terrain particulier, les ethnographes doivent aussi déterminer la meilleure façon de délimiter leur terrain. Aux méthodes applicables à des terrains situés exclusivement en ligne se sont ajoutées de nouvelles approches destinées à des terrains plus complexes d'un point de vue spatial, conçues dans le but d'analyser les liens complexes et contingents entre les espaces sociaux en ligne et hors-ligne. Les enjeux méthodologiques relatifs à la contribution que peut apporter à l'ethnographie l'expérience incarnée de la personne qui mène une enquête sur le terrain et les défis qui se profilent à l'horizon pour les ethnographes qui s'intéressent à l'exploration des terrains en ligne dans un contexte marqué par le développement de nouvelles plateformes d'interaction et par la présence toujours plus prégnante d'Internet dans nos vies sont toutes des questions abordées dans le présent chapitre.

L'ethnographie en ligne et l'ethnographie virtuelle

Un vaste éventail d'approches destinées à la recherche ethnographique en ligne ont émergé au fil des années dans différentes disciplines; elles portent des appellations variées, dont « ethnographie en ligne », « ethnographie virtuelle », « netnographie » ou « cyberethnographie ». L'étude de Baym (1995; 2000) à propos d'un groupe de discussion sur les téléromans a ouvert la voie à l'idée que les espaces virtuels pouvaient constituer des terrains d'enquête ethnographique. Même à cette époque, certains ethnographes ont observé que la notion de « présence en ligne » ne signifiait pas la même chose pour tout le monde. Parmi ceux-ci, Markham (1998) a préconisé une posture réflexive, ouverte aux différentes interprétations des internautes sur le sens de leurs activités en ligne, ainsi qu'aux émotions variées que suscite le fait d'habiter un espace en ligne. De façon similaire, j'ai décrit, en 2000, une ethnographie virtuelle qui considère les espaces en ligne comme des sites d'interaction à prendre au sérieux, sans aller toutefois jusqu'à soutenir qu'il existe un monde virtuel qu'on peut nettement distinguer du monde « réel ». Les approches ethnographiques ont évolué avec le développement d'Internet et la multiplication des contextes numériques. Kendall (2002) a réalisé son enquête de terrain sur le genre en participant à un forum de discussion et Senft (2008) a décrit grâce à des techniques d'observation et de participation ethnographiques l'expé-

rience des *camgirls* qui diffusent par webcam leur vie en direct sur Internet. La netnographie (Kozinets, 2009) a été élaborée pour faciliter l'étude des espaces en ligne, et elle est souvent utilisée en marketing pour sonder les motivations et les comportements de consommation. Parmi les ethnographies virtuelles à visée anthropologique, on compte l'étude du monde virtuel de *Second Life* réalisée par Boellstorff (2008), l'exploration du jeu *World of Warcraft* par Nardi (2010) et l'étude de Coleman (2013) sur une communauté de hackers, laquelle a exigé un important travail de terrain en ligne. Avec le développement du réseau Internet sont ainsi apparues différentes formes d'ethnographie en ligne fondées sur les principes de la méthode mère, mais que les praticiennes et praticiens de la recherche ont adaptées aux caractéristiques des interactions en ligne et aux thématiques privilégiées par leurs disciplines d'appartenance.

La nature et la finalité des recherches ethnographiques consacrées aux espaces en ligne ont évolué avec le temps. Selon Robinson et Schulz (2009), l'ethnographie en ligne s'est développée en trois phases distinctes. Sont apparues, dans un premier temps, des approches pionnières qui ont vu dans Internet un nouvel espace de construction de l'identité et insisté sur le caractère distinct des formations sociales en ligne; dans un deuxième temps, des approches de légitimation axées sur le transfert des préoccupations méthodologiques de l'ethnographie classique aux terrains en ligne, empreintes d'un certain scepticisme à l'égard du caractère distinct de ce dernier; enfin, plus récemment, des approches multimodales qui incorporent, en plus des données textuelles, des données audio et vidéo, et s'efforcent de contextualiser les interactions en ligne par rapport aux espaces hors-ligne. Robinson et Schulz (2009) avancent que les méthodes se sont affinées au fil des transformations du Web et de l'évolution des ambitions en matière de recherche. Les ethnographes entretiennent différentes conceptions sur la nature de l'espace en ligne, et leurs affiliations disciplinaires, leurs visées théoriques et leurs influences méthodologiques varient. L'ethnographie en ligne n'a pas de modèle unique; elle est plurielle, regroupant une diversité d'approches faisant d'un espace en ligne l'objet ou le terrain de l'enquête.

Le champ des études sur Internet (*Internet studies*) offre un terrain fertile aux développements méthodologiques, que ce soit pour l'ethnographie en ligne ou dans d'autres domaines. L'évolution des méthodes accompagne celle du Web, particulièrement depuis l'apparition des médias

sociaux. Dans leur recension des méthodes de recherche sur les médias sociaux, Giglietto *et al.* (2012) relèvent que les courants méthodologiques prédominants varient selon les plateformes. Ils répartissent ces courants en trois groupes : les méthodes ethnographiques, les méthodes statistiques et les méthodes computationnelles. Même si les premières ont donné le ton sur bon nombre de terrains en ligne, ce sont les méthodes computationnelles qui ont eu la préséance dans le cas de Twitter et ce n'est qu'ultérieurement qu'elles se sont vu complétées par des approches qualitatives. On peut affirmer, par conséquent, que les médias sociaux ont radicalement transformé le paysage de la recherche en sciences sociales, en raison notamment de l'accès qu'ils permettent à de vastes ensembles de données sur les activités quotidiennes, ce qui a favorisé en retour le développement des méthodes computationnelles. En plus d'adapter les méthodes existantes aux nouvelles conditions, certains ont mis au point des méthodes numériques novatrices spécialement destinées à la recherche en sciences sociales (Marres, 2012 ; Orton-Johnson et Prior, 2013 ; Rogers, 2013 ; Ruppert *et al.*, 2013). Malgré tout, l'avènement des « données massives » ne signifie pas pour autant qu'il faille délaisser les méthodes qualitatives comme l'ethnographie. Leur apparition aura favorisé l'analyse computationnelle, certes, mais d'aucuns insistent sur la nécessité de continuer également à privilégier des approches qualitatives à petite échelle, en faisant valoir qu'elles permettent de découvrir comment se dessinent les grandes tendances et d'analyser le sens que leur prêtent les individus participant à la recherche (boyd et Crawford, 2012). Dans le champ des études sur Internet, des méthodologies mixtes de recherche ont émergé qui offrent la possibilité de conjuguer des perspectives à grande et à petite échelle et d'étudier à la fois les *patterns* d'usage et leurs significations (Hesse-Biber et Griffin, 2013). Plus loin dans ce chapitre, j'avancerai que les méthodes ethnographiques qui permettent de poser à la fois un regard extérieur sur l'ancrage des activités dans différents contextes et un regard intérieur sur l'expérience relationnelle qu'elles procurent (grâce aux récits autoethnographiques) peuvent s'avérer des plus utiles dans la compréhension des formes de sociabilité que favorise le Web tout en offrant un complément approprié, ou même un contrepoids, aux approches fondées sur les données massives.

Ainsi, l'ethnographie joue depuis un certain temps un rôle notable dans la réflexion sur les conditions sociales propres aux environnements numériques. Elle conserve toute sa pertinence malgré l'avènement des

données massives et de l'analyse computationnelle à vaste échelle générée par les médias sociaux. Selon Coleman (2010), l'ethnographie des espaces en ligne revêt une importance particulière, puisque ceux-ci sont devenus des lieux d'expérience incontournables dans de nombreuses sphères de la vie quotidienne. Les travaux ethnographiques permettent de rendre compte de la place qu'occupent ces espaces et, toujours selon Coleman (2010), de contrer la tendance à voir le Web comme un tout homogène en faisant ressortir plutôt l'hétérogénéité et la spécificité des espaces en ligne. Alors que s'affirmait leur expérience, les ethnographes du Web y ont puisé une riche réflexion méthodologique, décortiquant les méthodes d'observation participante en ligne et se demandant dans quelle mesure les problèmes rencontrés étaient propres aux contextes numériques. La question de savoir si la pratique ethnographique en ligne, qu'on la désigne par les vocables « virtuelle », « cyber » ou « numérique », se distingue suffisamment de l'ethnographie traditionnelle pour mériter une démarcation nette, est une préoccupation récurrente dans les discussions. Elle sera abordée en détail dans la prochaine section.

L'ethnographie en ligne : méthodes et éthique

Les spécificités de la pratique ethnographique en ligne ont fait couler beaucoup d'encre. On relève dans les études une volonté d'expliquer comment les ethnographes du Web doivent procéder, préoccupation qui ne s'observe pas au même degré dans les comptes rendus issus de terrains plus classiques. Dans une certaine mesure, il s'agirait plutôt d'une question d'adaptation à des problèmes de nature technique. D'abord, les formes de présentation de soi et d'interaction qui s'offrent à la chercheuse et au chercheur varient selon les différentes plateformes. Ensuite, les archives numériques permettent de remonter dans le temps, un potentiel que les ethnographes continuent d'explorer. Enfin, il y a la nécessité de s'adapter aux formats multiples que prendront les données à recueillir et à analyser. Ce sont toutes là des questions pratiques concernant les défis qui accompagnent n'importe quel terrain d'enquête et les décisions à prendre une fois sur place. Il est donc très avantageux de pouvoir profiter de l'expérience d'autres ethnographes ayant enquêté sur des terrains semblables aux nôtres. On trouve aujourd'hui un ouvrage très utile qui traite de ces questions et présente un survol des techniques et des approches employées par

les ethnographes dans les mondes virtuels (Boellstorff *et al.*, 2012). Cela dit, en pratique, cet ouvrage – tout comme la majorité des discussions autour de l’ethnographie en ligne – aborde également des questionnements fondamentaux sur la manière de signaler sa présence, mais aussi sur les rôles à endosser, sur la façon de procéder pour la collecte des données et sur le statut à accorder aux connaissances issues de ce type d’enquête. Comme l’expliquent Robinson et Schultz (2009), le désir de donner à l’ethnographie en ligne ses lettres de noblesse, au même titre que l’ethnographie classique, a motivé la publication d’un nombre appréciable d’articles sur le sujet.

Steinmetz (2012) relève que l’identité et l’authenticité font partie des préoccupations récurrentes chez les ethnographes en ligne. Certains dilemmes se posent en effet depuis les tout débuts (Paccagnella, 1997). Comment évaluer l’authenticité dans un contexte virtuel? Peut-on se contenter de notions purement digitales de l’authenticité et de l’identité ou faut-il trianguler les observations faites en ligne avec d’autres sources de données? Dans toute enquête en ligne se pose la question du rôle que doit assumer l’ethnographe et de l’incidence qu’aura cette posture sur sa connaissance du milieu. La majorité des recherches sociologiques réalisées sur des terrains en ligne et qualifiées d’ethnographiques impliquent bel et bien une forme de participation. Par contre, certains de ces travaux poussent le concept d’ethnographie bien au-delà de la tradition des enquêtes immersives et expérientielles. À titre d’exemple, Frederick et Perrone (2014) qualifient d’ethnographie leur enquête sur les petites annonces affichées dans Craigslist, parce qu’ils disent avoir recouru à des méthodes ethnographiques d’analyse de contenu (Altheide, 1987) et à une forme d’ethnographie dite « instantanée » (Ferrell *et al.*, 2008). Comme le font valoir Garcia *et al.* (2009), se contenter d’observer un site (*lurk* en anglais) ou d’extraire des données sans s’immerger soi-même dans le milieu à l’étude pose certains problèmes sur le plan de l’interprétation ethnographique, alors qu’une participation active offre un avantage épistémique considérable. Toutefois, le fait de vanter les mérites de celle-ci ne résout pas certains problèmes, comme la forme précise qu’elle devrait prendre et la mesure dans laquelle l’ethnographe devrait jouer un rôle comparable à celui des autres usagers et usagères. Le débat se poursuit autour de la question de la coprésence entre l’ethnographe du Web et les personnes informatrices (Bengtsson, 2014).

Les préoccupations d'ordre éthique reviennent constamment dans les commentaires sur l'ethnographie virtuelle. Pour Robinson et Schulz (2009), l'une des principales questions à résoudre consiste à décider s'il existe des manières adéquates pour l'ethnographe d'informer les participants et participantes de sa présence selon les plateformes. Murthy (2008) s'interroge sur la tendance à mener des ethnographies en ligne à leur insu, qu'il juge troublante. Garcia *et al.* (2009) abordent pour leur part un certain nombre de dilemmes, dont la nécessité de déterminer si les espaces en ligne relèvent de la sphère publique, si l'ethnographe doit dévoiler son identité (et de quelle façon) et comment protéger la vie privée et l'autonomie des personnes à l'étude. On peut avancer que ce sont là des enjeux éthiques qui ne diffèrent pas des questions que se posent sur les terrains hors-ligne les ethnographes, qui doivent aussi composer avec les notions de sphère publique et de vie privée et décider de leur propre degré de visibilité auprès des sujets de leur enquête. Si la question éthique fait débat, c'est sans doute à cause de cette tendance observée chez quelques ethnographes qui considèrent les espaces en ligne comme des lieux où les règles d'éthique ne s'appliquent pas, plutôt qu'en raison d'une différence fondamentale entre espaces en ligne et hors-ligne. En un certain sens, l'ethnographie en ligne en tant que champ de recherche distinct a contribué à faire ressortir la nécessité de jeter un regard neuf sur les questions d'éthique. Markham (2006) argumente en faveur d'une approche située suivant laquelle chaque terrain ferait l'objet d'un questionnement particulier, au lieu de devoir satisfaire à un ensemble de règles préétablies convenant à tous les cas de figure. À cet égard, soulignons que l'Association of Internet Researchers propose une liste de ressources utiles précisément fondées sur l'approche située en matière d'éthique de la recherche en ligne (Ess et AoIR Ethics Working Committee, 2002; Markham et Buchanan, 2012).

Plus récemment, on s'est demandé dans quelle mesure les nouvelles plateformes viennent modifier la façon dont les problèmes méthodologiques se posent en ethnographie. Underberg et Zorn (2013) explorent les possibilités que présentent les technologies numériques en matière de *design* participatif, utilisant celles-ci pour construire des représentations culturelles avec et pour les internautes. Dans la même veine, Gallagher *et al.* (2013) soutiennent que les médias sociaux favorisent les projets collaboratifs fondés sur la participation active des individus à la démarche scientifique. Baker (2013) montre que les médias sociaux peuvent constituer à la fois une source d'information,

un outil pour rester en contact avec les usagers et usagères et un modèle de contextualisation d'autres types de données. Murthy (2013) réfléchit aux possibilités offertes, dans le champ des études organisationnelles, par les projets qui font appel aux réseaux sociaux et aux téléphones intelligents. L'évolution des technologies participatives a des implications pour la conception qu'on se fait du terrain d'enquête: selon Postill et Pink (2012), l'avènement du Web 2.0 a soulevé de nouveaux sujets de préoccupation et incité au dépassement des notions de communauté et de réseau, lesquelles auraient à leur avis accaparé inutilement la réflexion ethnographique sur les activités en ligne. On peut ainsi résumer les principales questions d'ordre méthodologique qui se posent en matière d'ethnographie virtuelle:

- Comment juger de l'authenticité et de la pertinence de trianguler les observations en ligne avec d'autres modalités d'observation ?
- Faut-il participer activement en ligne et si oui, de quelle manière ?
- Quels moyens faut-il prendre pour respecter les principes d'éthique, ainsi que la vie privée et l'autonomie des personnes ?
- Comment s'adapter aux dispositifs émergents qui ne cadrent pas bien avec ce que l'on entend par communauté en ligne ?

Même si ces enjeux sont récurrents dans bon nombre d'études menées en ligne, les solutions dépendent largement du contexte et des visées théoriques de la recherche. Les résultats d'une grande partie des travaux ethnographiques sur les espaces en ligne démontrent précisément qu'il ne faut pas s'attendre à appliquer les mêmes stratégies méthodologiques à toutes les plateformes. Par exemple, les problèmes qui se poseront seront très différents selon que le terrain est constitué par une liste de discussion, *Second Life*, WhatsApp ou Facebook; malgré tout, certains dilemmes seront en tout point semblables à ceux qui s'observent sur des terrains traditionnels. Quand l'ethnographe se tourne vers le Web, ses techniques peuvent différer de celles qui seraient mobilisées sur un terrain traditionnel, ce qui peut faire surgir des questions inédites sur le plan éthique. Néanmoins, nous avons tout à gagner à nous rappeler que le travail ethnographique et les difficultés qui l'accompagnent sont à bien des égards les mêmes en ligne et hors-ligne (Marshall, 2010).

Par conséquent, la définition d'un ensemble d'enjeux propres à l'ethnographie en ligne n'est pas toujours utile, même si le fait d'en déli-

imiter les contours permet d'attirer l'attention sur des méthodes potentiellement fécondes et de baliser un corpus de méthodes susceptible de lui conférer une certaine légitimité (Kozinets, 2012). La volonté de circonscrire l'ethnographie en ligne signale parfois une visée théorique particulière, à savoir celle de définir ce qui caractérise cette culture numérique désormais si présente dans la vie contemporaine (Boellstorff, 2010 ; Miller et Horst, 2012). Toutefois, ce n'est pas là le but premier dans bon nombre de recherches en sciences sociales. Leurs préoccupations théoriques peuvent les attirer vers une forme d'ethnographie comportant des activités en ligne, sans que le numérique constitue pour autant un thème central de leur réflexion. Dans la partie suivante, nous passons en revue un éventail de méthodes de recherche applicables à différents terrains en ligne, mais qui ne s'y cantonnent pas. Comme nous le verrons, la définition du terrain d'enquête revêt ici une importance toute particulière.

Le terrain d'enquête : approches combinée, multisite, en réseau et connective

Tout ethnographe doit posséder une connaissance pratique du terrain sur lequel portera son enquête. Cette exigence peut poser certaines difficultés dans les situations où une partie de l'enquête se déroule en ligne et l'autre partie hors-ligne, comme cela se produit fréquemment aujourd'hui (Hine, 2008 ; Steinmetz, 2012 ; Tuncalp, 2014). L'intégration d'Internet sous plusieurs formes au cœur de notre quotidien vient complexifier l'idée d'un terrain ethnographique qui serait préexistant et parfaitement circonscrit. La téléphonie et le réseau Internet mobiles ajoutent à la difficulté de délimiter le terrain d'enquête, car ces dispositifs sont présents dans toutes les sphères et désormais perçus comme allant de soi (Ling, 2012). Dans ce contexte, il devient de plus en plus ardu de cerner « les circonstances variablement dynamiques et changeantes » de la vie des individus (Gold, 1997, p. 395). L'ethnographe se retrouve invariablement face à un dilemme au moment de prendre des décisions à ce sujet. La reconnaissance du caractère contingent du terrain n'est pas propre à l'ethnographie du Web, la nature construite du terrain étant déjà admise en anthropologie (Gupta et Ferguson, 1997 ; Amit, 1999), et en sociologie dans une moindre mesure. Malgré cela, l'intérêt grandissant des sciences sociales pour les activités

ayant lieu sur des terrains médiatisés dynamiques et fragmentés nous fait prendre nouvellement conscience de ces enjeux.

Garcia *et al.* (2009) soutiennent fermement que peu de questions trouveront de réponse par la seule exploration du Web, même s'il est difficile d'éviter d'en tenir compte d'une manière ou d'une autre dans l'étude de la société contemporaine. Dans un même ordre d'idées, Hallett et Barber (2014) affirment que l'étude du mode de vie de nombreux groupes sociaux doit désormais prendre au sérieux à la fois leurs activités en ligne et hors-ligne. Une position qui diverge, jusqu'à un certain point, de celle que défend Boellstorff (2010) dans une discussion sur les différentes notions de terrain, où il vante la pertinence des terrains exclusivement en ligne. Néanmoins, ces travaux se rejoignent sur un point : la définition du terrain doit convenir à la question de recherche. J'estime (2015) que l'effort de définition d'une ethnographie spécifiquement en ligne voit son utilité diminuer face à un Internet imbriqué dans une multitude de contextes où l'on construit du sens, tant en ligne que hors-ligne. Les méthodes fondées sur une approche combinée (*blended*), mobile ou en réseau peuvent aboutir à des résultats très profitables, car elles permettent aux ethnographes d'illustrer la complexité de l'expérience vécue dans différents espaces, sans être contraints de se limiter à une simple description du terrain en ligne qui fait l'objet de son enquête. Les réflexions théoriques menées dans d'importants champs de recherche en sociologie, sur la nature et la portée des mouvements sociaux, par exemple, ou sur la construction de l'identité, se cantonnent rarement à comprendre seulement l'activité des internautes au sein d'un média particulier, d'où la nécessité d'une démarche couvrant plusieurs médias à la fois. De même, les enjeux liés à l'élaboration de politiques prennent rarement pour seul objet les activités se déroulant en ligne ; ces politiques s'intéressent également à l'expérience et aux pratiques des individus dans d'autres sphères. Dyke (2013) a conjugué, dans le cadre d'une approche combinée, une enquête auprès d'une communauté virtuelle axée sur la promotion de l'anorexie à un programme de prévention des troubles alimentaires déployé dans des écoles et des maisons de jeunes, s'intéressant aux déplacements des jeunes entre les espaces en ligne et hors-ligne. Son étude lui a permis de formuler des recommandations qui font ressortir la nécessité, pour qui veut mettre en place des interventions réussies, de comprendre la façon dont les jeunes risquant de souffrir de troubles alimentaires naviguent d'un contexte à l'autre. Malgré les diffi-

cultés que peut présenter ce genre de démarche, ces défis peuvent être éclairants tant sur le plan théorique que sur le plan pratique.

La recherche sur les mouvements sociaux semble particulièrement tirer profit des approches multimodales et multisite. Postill et Pink (2012) préconisent une analyse du réseau « désordonné » d'interconnexions entre les médias sociaux qui acquièrent une importance particulière aux yeux des groupes qui les utilisent. À leur tour, Treré (2012) ainsi que Farinosi et Treré (2011) prônent la nécessité de mener des études multimodales qui tiennent compte des médias sociaux en tant que sites où surviennent des événements, tout en contextualisant cette activité par rapport aux situations concrètes que l'action militante aspire à transformer. Outre l'étude des mouvements sociaux, les recherches ethnographiques fondées sur une approche multisite ou en réseau comprennent l'étude sur le poker de Farnsworth et Austrin (2010); l'enquête menée par Burrell (2009) sur l'usage d'Internet à Accra; l'étude de Beneito-Montagut (2011) sur les émotions en ligne et l'enquête multisite d'Orton-Johnson (2014) sur le tricot et les espaces virtuels. Bien que la netnographie se pratique bien souvent uniquement en ligne, Kozinets (2009) évoque la possibilité d'associer cette méthode à l'ethnographie traditionnelle. Parmi les études qui se réclament explicitement d'une approche associant l'ethnographie et la netnographie figure l'enquête réalisée par Nichols et Rine (2012) sur les identités narratives au sein d'une communauté désindustrialisée. Je soutiens que la méthode connective est utile à l'ethnographie d'un Internet « imbriqué, incarné et quotidien² ». Cette approche tient compte des cadres multiples dans lesquels s'inscrit la construction du sens des activités en ligne et considère comme des choix stratégiques les liens que l'ethnographe décide d'étudier, plutôt que des choix dictés par les frontières d'un terrain circonscrit *a priori*. Les approches qui combinent la recherche en ligne et hors-ligne offrent à l'ethnographe un éventail de moyens lui permettant de définir des points de départ et d'orienter sa démarche. On pourra par exemple commencer par l'observation en ligne, puis mener des entrevues hors-ligne qui serviront à contextualiser les phénomènes observés, comme l'a fait Sade-Beck (2008) dans une enquête sur des communautés de deuil. Une étude combinée peut aussi démarrer hors-ligne, font valoir Miller et Slater (2000). Il est même possible de mener des

2. Internet « E au cube » en anglais : *embedded, embodied, everyday* (NDT).

recherches sur des phénomènes en ligne en adoptant principalement une perspective hors-ligne. Mabweazara (2010) a ainsi étudié l'usage que font les journalistes des technologies numériques en s'appuyant sur une approche ethnographique traditionnelle. Commentant la tâche difficile qu'a représentée l'analyse d'un mode de travail aussi dispersé et fragmenté, il avance que l'observation participante et l'entrevue en profondeur permettent un point de vue différent sur les usages de ces technologies. Selon lui, il est difficile, mais non impossible, d'étudier l'imbrication des activités en ligne dans les contextes hors-ligne en suivant une démarche ethnographique. Le choix du *design* d'enquête, en ce qui touche notamment la décision de recueillir des données hors-ligne ou en ligne, ou le choix de se concentrer sur la communication synchrone ou asynchrone, dépendra des aspirations de l'ethnographe et des caractéristiques du groupe à l'étude (Wilkerson *et al.*, 2014).

Les terrains qui incluent des activités en ligne peuvent prendre de multiples formes. Les raisons motivant le genre d'activités à observer ou auxquelles participer varieront en fonction des objectifs poursuivis. Voici un aperçu de quelques approches utiles à la définition d'un terrain, variant selon la nature des activités concernées ou de la stratégie employée par l'ethnographe pour se déplacer entre les sites (même si les termes employés pour désigner ces approches varient selon les écrits et que les chevauchements sont fréquents) :

- *En ligne* : l'ethnographe limite son enquête aux activités qui se déroulent au sein d'un espace en ligne donné (ou d'un ensemble d'espaces interconnectés), sans chercher à les situer dans un cadre hors-ligne.
- *Multimodale* : l'ethnographe analyse tous les modes de communication utilisés par le groupe qui fait l'objet de ses recherches (communication en face à face, documents, téléphone, médias sociaux, autres espaces en ligne, etc.).
- *Multisite* : l'ethnographe sélectionne, avant d'entamer ses recherches ou à mesure qu'elles progressent, un ensemble de sites interreliés qui l'aideront à éclairer les différentes facettes de l'expérience qui l'intéresse.
- *Combinée* : l'ethnographe conjugue deux approches ou davantage – souvent l'observation en ligne et hors-ligne – dans le but d'étudier un phénomène donné. La combinaison des méthodes choisies peut

varier ; il pourra s'agir d'une comparaison structurée entre les deux méthodes retenues ou bien d'un jumelage plus dynamique impliquant des déplacements entre les sites, à l'instar des approches en réseau, multisite et connective.

- *En réseau* : l'ethnographe parcourt un réseau d'interconnexions en suivant le flux des échanges communicationnels entre les membres d'un groupe ou autour d'une activité donnée, ce qui implique qu'on peut découvrir en cours de route des sites dont on n'avait pas anticipé l'existence.
- *Connective* : l'ethnographe se déplace entre différents modes de communication et différents lieux (en ligne ou hors-ligne). On oriente ses choix en fonction de préoccupations théoriques axées sur les rapports contingents qui émergent à mesure que les sujets s'approprient les activités en ligne et leur prêtent une signification dans un contexte hors-ligne et vice versa.

Comme le font observer Postill et Pink (2012), il peut s'avérer utile, dans maintes situations, de laisser ouverte la question touchant la nature du lien entre les activités qui se déroulent sur les médias sociaux et dans des lieux d'interaction en face à face. En effet, l'ethnographe doit s'appliquer à découvrir les formes de sociabilité qui se déploient à la fois en ligne et hors-ligne, et ce, dans l'optique plus large de cerner la réalité sociopolitique d'activités mettant en jeu des médias sociaux et des espaces en ligne, sans toutefois s'y cantonner. Dans ces domaines d'activités médiatisés et fragmentés, la construction du terrain s'accomplit de manière dynamique. Par conséquent, la multitude de possibilités qui s'offrent à l'ethnographe implique une responsabilité accrue et exige une démarche réflexive quant aux décisions à prendre pour la construction du terrain.

L'autoethnographie en ligne

La question de la réflexivité est un élément essentiel dans la définition du terrain, mais son importance ne s'arrête pas là : elle concerne aussi les recherches impliquant des activités en ligne. On assiste en ethnographie en ligne à l'émergence d'un nouveau courant de pensée qui considère l'expérience incarnée de la chercheuse et du chercheur comme une source de connaissance valable. Les études qui s'en réclament voient dans l'auto-

ethnographie un moyen de tirer profit de la subjectivité et de mettre au jour des dynamiques affectives qui se prêtent moins bien aux analyses traditionnelles et qui sont limitées, dans une certaine mesure, par ce que les personnes participantes parviennent à exprimer et à communiquer à l'ethnographe (Reed-Danahay, 1997; Ellis, 2004; Ellis *et al.*, 2010). Les autoethnographes sont souvent partie prenante des situations décrites dans leurs travaux. Quoiqu'elles demeurent rattachées à leur champ disciplinaire, restant ainsi au fait des thèmes potentiellement pertinents ou d'actualité au regard de divers courants théoriques, ces études mettent l'accent sur l'évocation et sont souvent peu étoffées sur le plan théorique ou bibliographique. Elles sont parfois accusées de complaisance, se concentrant trop souvent sur le vécu personnel au détriment d'autres points de vue (Sparkes, 2002). L'autoethnographie offre néanmoins une piste de solution aux défis que présente l'univers complexe, fragmenté et désordonné des connexions contingentes entre les espaces en ligne et hors-ligne. Elle permet notamment une réflexion sur le ressenti de l'ethnographe qui expérimente ces connexions.

Par ailleurs, le terme « autoethnographie » renvoie à des situations courantes où les personnes décrivent et commentent leur propre situation, plutôt qu'à un projet ethnographique particulier. L'autoethnographie peut s'avérer utile à l'étude des interactions en ligne. Par exemple, les commentaires publiés sont traités comme des documents réflexifs qui fournissent une lecture de la culture dans laquelle les personnes évoluent (Nemeth et Gropper, 2008; Jacobs, 2010). Cet usage reste toutefois minoritaire; c'est une conception plus étroite de l'autoethnographie comme méthodologie délibérée qui semble plutôt avoir gagné du terrain au sein de la discipline. Appliquée aux espaces numériques, cette approche pourrait grandement contribuer à améliorer notre capacité à interpréter le sens que prêtent les internautes à leurs interactions.

D'une certaine manière, l'ethnographe du Web a toujours une posture participante, ne serait-ce que parce que sa présence en ligne l'oblige à utiliser les mêmes technologies que les autres internautes. Pour analyser un site de jeux vidéo, par exemple, il faudra peut-être acquérir les habiletés requises pour rester dans le jeu assez longtemps pour bien comprendre le contexte, avant même d'envisager ce qu'on peut retirer de l'expérience immersive du point de vue épistémologique. Sur le continuum allant de l'observation participante à la participation observante (Gold, 1958),

l'ethnologue du Web restera toujours, dans une certaine mesure, en posture d'observation participante, de par son utilisation – pour interagir avec son terrain – du même dispositif que les personnes observées même si, comme ces dernières le font souvent, l'ethnologue peut avoir recours à une panoplie de modes de communication. L'analyse des conditions sociales produites par l'usage d'un dispositif de communication particulier peut fournir des éléments de compréhension importants à l'ethnologue du Web.

Markham (1998) est l'une des pionnières de l'approche autoethnographique appliquée au contexte en ligne, ayant cherché à décrire l'expérience que procurent la navigation et la communication en ligne sous différentes formes. Son analyse met en lumière le caractère contingent de cette expérience. Elle rapporte qu'elle-même et les personnes rencontrées en ligne n'envisageaient pas toujours la communication médiatisée par ordinateur en termes de lieu, mais pouvaient la percevoir aussi comme un outil ou une manière d'être. D'autres autoethnographies de l'expérience en ligne ont été réalisées à la suite de Markham, ce type de démarche jouissant d'une légitimité de plus en plus grande, même si on la considère généralement comme une approche alternative ou « de niche ». Cette méthode permet effectivement de cerner de façon nuancée les pratiques identitaires en ligne. Par exemple, Dumitrica et Gaden (2008) en ont fait un puissant outil d'analyse de l'expérience du genre dans l'univers virtuel de *Second Life*. Gatson (2011) a réalisé une étude sur un « projet d'autoconstruction » (*selfing*³) au moyen d'une participation aux activités en ligne et hors-ligne, d'une enquête systématique et d'une réflexion sur ses propres pratiques en la matière, démarche qui se rapproche de l'autoethnographie même si elle ne s'en réclame pas.

Les autoethnographies se sont avérées efficaces dans les études portant sur la cyberpédagogie. Lee (2008) y a recours pour décrire son expérience d'enseignement dans le cadre d'un cours en ligne. Comme plusieurs exemples d'écriture autoethnographique, son récit relève davantage de l'évocation et n'apporte pas de contribution théorique particulière. Elle y relate comment les activités relevant de la modération d'un cours en ligne se mêlent aux pressions et aux sensations vécues dans son quotidien hors-ligne. Son ambition théorique reste implicite: elle nous décrit les

3. Le terme a été corrigé par l'autrice par rapport à l'édition originale.

détails de son expérience sans en tirer des conclusions qui s'inscriraient dans un cadre d'analyse sociologique traditionnel. Une autre étude autoethnographique, celle de Kruse (2013), s'intéresse à l'expérience étudiante, par contraste avec la position de tuteur, dans le cadre d'un cours en ligne axé sur l'apprentissage de la mandoline. Ce mode d'apprentissage a suscité chez lui une réflexion sur l'établissement des liens dans un tel contexte et sur le sentiment d'isolement qu'il a éprouvé. Henning (2012) relate pour sa part l'expérience d'une enseignante devenue étudiante à son tour. Ayant elle-même donné des cours en ligne, elle raconte ce qu'elle a découvert en endossant pour la première fois ce rôle. Dans un autre exemple, Tschida et Sevier (2013) se fondent sur une approche autoethnographique pour analyser le déroulement d'un cours proposé en ligne, faisant ressortir les difficultés qui se sont présentées sur le plan pédagogique en matière de pratiques et d'attentes, par comparaison avec l'enseignement en face à face.

Une autre approche est celle de l'autonethnographie, qui se définit comme une forme d'observation fondée sur la recherche, l'observation non participante (ou *lurking*) et la prise de notes réflexive (Kozinets et Kedzior, 2009). Wilkinson et Patterson (2014) ont analysé des *mashups* de la série *Peppa Pig* créés par des internautes et diffusés sur YouTube, puis ont complété leur analyse par des entrevues en ligne sollicitées par vidéo. Le terme « autonethnographie » renvoie ici à une période d'observation préalable axée sur l'expérience de l'ethnographe en tant qu'agent actif dans l'exploration d'un territoire, puis à l'examen des habitudes de consommation en ligne, comme c'est le cas des nombreuses études qui s'inscrivent dans ce courant. Dans une démarche similaire, et sans toutefois qualifier leur étude d'ethnographique ou de netnographique, Beer et Penfold-Mounce (2009) se sont intéressés aux potins sur les célébrités qui circulent sur Internet, en adoptant une position semblable à celle de tout autre internaute en quête de ce genre d'informations.

Dans d'autres cas, l'usage que fait l'ethnographe de sa propre expérience dépasse la simple réflexion sur le fait de s'immerger dans le même dispositif que les personnes à l'étude, pour s'étendre à la question de son rôle dans la situation. Ainsi, Baym (1995; 2000) a évolué du statut de participante à part entière dans un groupe de discussion sur les téléromans à celui d'ethnographe du groupe. Hughey (2008) a d'abord participé à un groupe consacré aux fraternités et aux sororités afro-américaines avant

d'en faire son objet d'étude, assumant dès lors un rôle qu'il décrit comme celui d'un participant observateur, qui combine l'observation participante active (incluant la création de fils de discussion), l'analyse de contenu et les entrevues. Dans les études portant sur le phénomène des admirateurs (*fan studies*), il arrive souvent que l'ethnographe soit déjà membre du groupe en question, ce qui expliquerait peut-être le nombre croissant d'autoethnographies portant sur ce type de communauté virtuelle. Monaco (2010) a ainsi fréquenté les sites et les activités en ligne des amateurs d'une émission de télévision en particulier, pour ensuite analyser l'évolution de son rapport au récit découlant de sa participation à la communauté. L'autoethnographie offre un moyen d'étudier les pratiques de navigation d'au moins un membre de l'auditoire et le sens qu'il accorde aux différentes manifestations en ligne et hors-ligne de l'objet d'attention de la part des fans. Le caractère aléatoire des relations qui s'établissent constituerait tout un défi sur le plan méthodologique pour l'ethnographe en situation d'observation, en raison de la difficulté à suivre les internautes d'un site à l'autre. Dans le cadre d'une étude autoethnographique sur l'univers des amateurs de football, Parry (2012) réfléchit au rôle que jouent les sites dédiés à ce sport dans l'expérience de ces personnes, au sens qu'on accorde à ce type d'engagement à mesure qu'on devient adepte, ainsi qu'aux incidences et aux ressentis que cet engagement génère dans le cadre d'autres activités quotidiennes. Enfin, Paechter (2013) porte à la fois un regard intérieur et extérieur sur un groupe dont elle était membre depuis un certain temps, menant ouvertement (mais rétrospectivement) une étude où elle analyse des messages archivés, certains remontant à l'époque de sa participation.

La perspective autoethnographique implique que le regard est tourné vers l'intérieur, ce qui permet d'éclairer une expérience selon un point de vue particulier. Toutefois, elle ne libère pas l'ethnographe de l'obligation d'élargir sa réflexion aux structures et aux contraintes sociales présentes dans la vie quotidienne des individus. Les personnes qui pratiquent l'autoethnographie se considèrent comme des êtres sociaux. Leurs réflexions sur leurs expériences en ligne nous aident à découvrir en quoi des infrastructures types peuvent donner lieu à des expériences personnelles et comment les modes d'interaction en ligne façonnent ce que nous pouvons devenir les uns pour les autres en tant qu'êtres humains. L'autoethnographe contextualise les expériences en ligne et analyse les

sens multiples qu'elles revêtent. Comme cette démarche procède d'une expérience subjective, elle peut offrir une perspective multidimensionnelle du Web qui n'est pas assujettie à une idée préconçue du terrain censé convenir au phénomène qu'elle veut étudier. Tout comme les approches combinée, multisite, en réseau et connective, l'autoethnographie est susceptible d'alimenter des visées théoriques ambitieuses, car elle permet de suivre la trace de phénomènes sans avoir à se limiter superficiellement à une définition de l'expression « en ligne » surdéterminée par les caractéristiques d'un dispositif donné.

La voie à suivre...

On peut affirmer qu'Internet s'est infiltré dans la société contemporaine au point de devenir invisible (Parks, 2009). Les gens se connectent aujourd'hui sans réfléchir; ils ne se limitent plus à un seul moyen de communication à la fois, mais les combinent en fonction de leurs besoins et pratiquent le multitâche d'un appareil à l'autre. Après avoir traversé une crise des représentations (Denzin, 1997), les ethnographes font aujourd'hui face à une nouvelle crise qui concerne cette fois l'agentivité. Nous nous trouvons face à une multitude de choix au moment de définir un terrain d'enquête et devons en outre assumer la pleine responsabilité de nos décisions, quelles que soient les circonstances. On ne peut plus s'en remettre à l'idée qu'il existerait des terrains prédéfinis qui n'attendent que l'arrivée d'un ethnographe. Le terrain doit désormais être circonscrit en fonction des visées du projet. Dans certains cas, il sera situé exclusivement en contexte numérique. Or les ethnographes devront souvent se déplacer au-delà de contextes entièrement en ligne et finalement délaissier toute conception *a priori* des frontières de leur terrain.

Outre la nécessité pour l'ethnographe de prendre conscience de son agentivité dans la construction de son terrain et de l'assumer, il faut aussi prendre en considération quelques-unes des formes d'agentivité moins visibles qui relèvent du terrain, car elles peuvent nous orienter dans une direction plutôt qu'une autre et ainsi modeler imperceptiblement nos recherches. Il nous faut réfléchir aux diverses formes de résistance que le terrain oppose à notre quête de connaissance. Parmi ces obstacles, mentionnons la commercialisation et la privatisation des espaces en ligne, auxquelles s'ajoute la relative invisibilité des pratiques de consommation

des contenus numériques. Du fait de leur commercialisation, certaines parties du Web sont moins accessibles à l'étude ethnographique (Kozinets, 2009), car les préoccupations relatives à la propriété des données l'ont transformé en une enfilade de mondes fermés et protégés par des mots de passe (Lievrouw, 2012). Cette situation peut amener l'ethnographe à négocier un consentement éclairé non seulement des internautes qui publient des contenus en ligne, mais aussi des compagnies, les internautes ne sachant parfois même pas que celles-ci revendiquent la propriété de leurs contenus. Les ethnographes ne peuvent pas tenir pour acquis que leurs informatrices et informateurs sont autorisés à donner accès à ce matériel ; du fait des droits de propriété, certains terrains en ligne pourraient bien rester hors de leur portée. Par ailleurs, les problèmes touchant la propriété des données et le droit à la vie privée incitent souvent les internautes à rechercher des formes d'interaction en ligne plus secrètes ou éphémères. Dans ce genre de cas, les ethnographes font face à la plupart des problèmes qui surgissent hors-ligne, notamment la difficulté de saisir la nature des interactions fugaces qui prennent place dans les espaces privés (Gehl, 2014). La promesse d'une vie sociale en ligne librement accessible à l'observation ethnographique pourrait s'avérer une situation passagère et relativement limitée. Il faut se garder de faire des généralisations en se fondant sur les facilités d'accès actuelles, mais fortuites, à certains phénomènes.

Le repli vers des formes d'interaction plus privées et éphémères soulève aussi une question plus large, à savoir celle des moyens à mettre en œuvre pour incorporer les pratiques relatives à la consommation et à l'interprétation dans l'enquête ethnographique. Bon nombre de personnes lisent du contenu en ligne sans publier elles-mêmes des messages. Leur activité ne laisse aucune trace observable, mais elle peut néanmoins être significative dans leur vie. L'ethnographie en ligne repose surtout sur les données fournies par des personnes contribuant activement, mais ces activités ne représentent en réalité que la pointe de l'iceberg, ce qui pose problème pour qui veut comprendre Internet dans sa globalité en tant que phénomène social. Pour analyser Internet en tant que phénomène social enraciné dans toutes les sphères de la vie contemporaine, il faut tenir compte des formes variées d'engagement en ligne, y compris leur incidence sur la manière dont les individus se pensent en tant qu'êtres sociaux et sur le développement de leur réflexivité quant à leur place dans le monde. Les

usages d'Internet qui revêtent une importance sociale ne laissent pas tous des traces durables ou facilement repérables au moyen des méthodes d'observation en ligne. Les études sur la consommation qui se fondent sur une approche combinée, multisite ou autoethnographique contribuent utilement à la compréhension de l'imbrication d'Internet dans la vie quotidienne. Il reste toutefois un travail considérable à faire avant de pouvoir élucider les conséquences de l'invisibilité des habitudes de consommation de contenus en ligne, laquelle entrave les efforts des ethnographes dont les travaux se concentrent sur la partie visible des pratiques numériques. Les approches anthropologiques de l'ethnographie se sont avérées particulièrement fécondes pour explorer les dimensions plus vastes d'une culture imprégnée par le numérique (Horst et Miller, 2012).

Comparativement à d'autres approches ethnographiques, la réalisation d'une enquête uniquement sur un terrain en ligne est relativement peu coûteuse et n'a souvent qu'une incidence limitée sur la vie de l'ethnographe. L'ethnographie en ligne a certainement une contribution essentielle à apporter, notamment en ce qui a trait à la dynamique des relations personnelles et aux pratiques contemporaines de construction de sens et de l'identité. Les mondes numériques procurent un accès sans précédent à la vie des individus et, par le fait même, offrent la possibilité de connaître et d'analyser leur quotidien en détail et en profondeur, ce qui devrait représenter un apport précieux pour les sciences sociales. Il est toutefois essentiel de préciser les visées théoriques qui guident nos travaux, car le fait de leur apposer l'étiquette « virtuelle » ou celle de « netnographie » peut parfois s'avérer contre-productif. Nos enquêtes de terrain doivent contribuer au développement des sciences sociales, qu'elles se déroulent exclusivement en ligne ou qu'elles se fondent sur une approche combinée impliquant des allers-retours entre des espaces médiatisés. Ces étiquettes ne sont pas forcément très utiles si elles ne constituent qu'un moyen de prendre ses distances par rapport à une sphère théorique plus large. Pour que nos contributions restent rigoureuses et pertinentes tant pour la production théorique que pour l'élaboration de politiques, il faudra entreprendre des enquêtes plus complexes et difficiles qui s'étendront de part et d'autre de la frontière en ligne/hors-ligne et qui s'intéresseront aux aspects moins visibles du numérique en tant que phénomène social. Il importe d'éviter de donner l'impression qu'il est suffisant de mener des enquêtes exclusivement en ligne. Dans le contexte hautement politisé du

financement de la recherche, où prédomine le souci de limiter les dépenses, il faut rappeler que le projet le moins coûteux n'est pas forcément le meilleur. Les recherches qualitatives complexes menées sur des terrains en ligne et hors-ligne ne sont ni rapides, ni faciles, ni économiques à réaliser. Toutefois, elles pourraient bien constituer l'une des meilleures façons de mettre au jour et d'analyser des formes de sociabilité émergentes qui résistent aux généralisations.

Références

- ALTHEIDE, David L., « Reflections: Ethnographic content analysis », *Qualitative Sociology*, vol. 10, n° 1, 1987, p. 65-77.
- AMIT, Vered (dir.), *Constructing the Field*, Routledge, 1999.
- BAKER, Sally, « Conceptualising the use of Facebook in ethnographic research: As tool, as data and as context », *Ethnography and Education*, vol. 8, n° 2, 2013, p. 131-145.
- BAYM, Nancy K., « The emergence of community in computer-mediated communication », dans Steve JONES (dir.), *Cybersociety*, SAGE Publications, 1995, p. 138-163.
- BAYM, Nancy K., *Tune In, Log On: Soaps, Fandom and Online Community*, SAGE Publications, 2000.
- BEER, David et Ruth PENFOLD-MOUNCE, « Celebrity gossip and the new melodramatic imagination », *Sociological Research Online*, vol. 14, n° 2, 2009, p. 1-12.
- BENEITO-MONTAGUT, Roser, « Ethnography goes online: Towards a user-centred methodology to research interpersonal communication on the Internet », *Qualitative Research*, vol. 11, n° 6, 2011, p. 716-735.
- BENGTSSON, Stina, « Faraway, so close! Proximity and distance in ethnography online », *Media, Culture and Society*, vol. 11, n° 6, 2014, p. 862-877.
- BOELLSTORFF, Tom, *Coming of Age in Second Life: An Anthropologist Explores the Virtually Human*, Princeton University Press, 2008.
- BOELLSTORFF, Tom, « A typology of ethnographic scales for virtual worlds », dans William SIMS BAINBRIDGE (dir.), *Online Worlds: Convergence of the Real and the Virtual*, Springer, 2010, p. 123-133.
- BOELLSTORFF, Tom, Bonnie NARDI, Celia PEARCE et T. L. TAYLOR, *Ethnography and Virtual Worlds: A Handbook of Method*, Princeton University Press, 2012.
- BOYD, danah et Kate CRAWFORD, « Critical questions for big data: Provocations for a cultural, technological and scholarly phenomenon », *Information, Communication & Society*, vol. 15, n° 5, 2012, p. 662-679.
- BURRELL, Jenna, « The field site as a network: A strategy for locating ethnographic research », *Field Methods*, vol. 21, n° 2, 2009, p. 181-199.
- COLEMAN, E. Gabriella, *Coding Freedom: The Ethics and Aesthetics of Hacking*, Princeton University Press, 2013.

- COLEMAN, E. Gabriella, « Ethnographic approaches to digital media », *Annual Review of Anthropology*, vol. 39, n° 1, 2010, p. 487-505.
- DENZIN, Norman K., *Interpretive Ethnography: Ethnographic Practices for the 21st Century*, SAGE Publications, 1997.
- DUMITRICA, Delia et Georgia GADEN, « Knee high-boots and six pack abs: Auto-ethnographic reflections on gender and technology in *Second Life* », *Journal of Virtual Worlds Research*, vol. 1, n° 3, 2008, p. 4-23.
- DYKE, Sarah, « Utilising a blended ethnographic approach to explore the online and offline lives of pro-ana community members », *Ethnography and Education*, vol. 8, n° 2, 2013, p. 146-161.
- ELLIS, Carolyn, *The Ethnographic I: A Methodological Novel about Autoethnography*, AltaMira Press, 2004.
- ELLIS, Carolyn, Tony E. ADAMS et Arthur P. BOCHNER, « Autoethnography: An overview », *Forum Qualitative Sozialforschung/Forum: Qualitative Social Research*, vol. 12, n° 1, art. 10, 2010.
- ESS, Charles et AOIR ETHICS WORKING COMMITTEE, *Ethical Decision-Making and Internet Research: Recommendations from the AoIR Ethics Working Committee*, Association of Internet Researchers, 2002.
- FARINOSI, Manuela et Emiliano TRERÉ, « Inside the “People of the Wheelbarrows”: Participation between online and offline dimension in the post-quake social movement », *Journal of Community Informatics*, vol. 6, n° 3, 2011.
- FARNSWORTH, John et Terry AUSTRIN, « The ethnography of new media worlds? Following the case of global poker », *New Media & Society*, vol. 12, n° 7, 2010, p. 1120-1136.
- FERRELL, Jeff, Keith J. HAYWARD et Jock YOUNG, *Cultural Criminology: An Invitation*, SAGE Publications, 2008.
- FREDERICK, Brian J. et Dina PERRONE, « “Party N Play” on the Internet: Subcultural formation, Craigslist, and escaping from stigma », *Deviant Behavior*, vol. 35, n° 11, 2014, p. 859-884.
- GALLAGHER, Kathleen, Anne WESSELS et Burçu Yaman NTELIOGLOU, « Becoming a networked public: Digital ethnography, youth and global research collectives », *Ethnography and Education*, vol. 8, n° 2, 2013, p. 177-193.
- GARCIA, Angela C. *et al.*, « Ethnographic approaches to the Internet and computer-mediated communication », *Journal of Contemporary Ethnography*, vol. 38, n° 1, 2009, p. 52-84.
- GATSON, Sarah N., « Self-naming practices on the Internet: Identity, authenticity, and community », *Cultural Studies, Critical Methodologies*, vol. 11, n° 3, 2011, p. 224-235.
- GEHL, Robert W., « Power/freedom on the dark web: A digital ethnography of the Dark Web Social Network », *New Media & Society*, 2014, vol. 18, n° 7, p. 1219-1235.
- GIGLIETTO, Fabio, Luca ROSSI et Davide BENNATO, « The open laboratory: Limits and possibilities of using Facebook, Twitter, and YouTube as a research data source », *Journal of Technology in Human Services*, vol. 30, n° 3-4, 2012, p. 145-159.
- GOLD, Raymond L., « The ethnographic method in sociology », *Qualitative Inquiry*, vol. 3, n° 4, 1997, p. 388-402.

- GOLD, Raymond L., « Roles in sociological field observations », *Social Forces*, vol. 36, n° 3, 1958, p. 217-223.
- GUPTA, Akhil et James FERGUSON, *Anthropological Locations: Boundaries and Grounds of a Field Science*, University of California Press, 1997.
- HALLETT, Ronald E. et Kristen BARBER, « Ethnographic research in a cyber era », *Journal of Contemporary Ethnography*, vol. 43, n° 3, 2014, p. 306-330.
- HENNING, Teresa Beth, « Writing professor as adult learner: An autoethnography of online professional development », *Journal of Asynchronous Learning Networks*, vol. 16, n° 2, 2012, p. 9-26.
- HESSE-BIBER, Sharlene et Amy J. GRIFFIN, « Internet-mediated technologies and mixed methods research: Problems and prospects », *Journal of Mixed Methods Research*, vol. 7, n° 1, 2013, p. 43-61.
- HINE, Christine, *Ethnography for the Internet: Embedded, Embodied and Everyday*, Berg Publishers, 2015.
- HINE, Christine, « How can qualitative Internet researchers define the boundaries of their projects? », dans Annette N. MARKHAM et Nancy BAYM (dir.), *Internet Inquiry: Conversations about Method*, SAGE Publications, 2008, p. 1-20.
- HINE, Christine, *Virtual Ethnography*, SAGE Publications, 2000.
- HORST, Heather A. et Daniel MILLER, *Digital Anthropology*, Berg Publishers, 2012.
- HUGHEY, Matthew W., « Virtual (br)others and (re)sisters: Authentic black fraternity and sorority identity on the Internet », *Journal of Contemporary Ethnography*, vol. 37, n° 5, 2008, p. 528-560.
- JACOBS, Katrien, « Lizzy Kinsey and the Adult Friendfinders: An ethnographic study of Internet sex and pornographic self-display in Hong Kong », *Culture, Health and Sexuality*, vol. 12, n° 6, 2010, p. 691-703.
- KENDALL, Lori, *Hanging Out in the Virtual Pub: Masculinities and Relationships Online*, University of California Press, 2002.
- KOZINETS, Robert V., « Marketing netnography: Promoting a new research method », *Methodological Innovations Online*, vol. 7, n° 1, 2012, p. 37-45.
- KOZINETS, Robert V., *Netnography: Doing Ethnographic Research Online*, SAGE Publications, 2009.
- KOZINETS, Robert V. et Richard KEDZIOR, « I, Avatar: Auto-netnographic research in virtual worlds », dans Nathalie T. WOOD et Michael R. SOLOMON (dir.), *Virtual Social Identity and Consumer Behavior*, ME Sharpe, 2009, p. 3-19.
- KRUSE, Nathan B., « Locating “The Road to Lisdoonvarna” via autoethnography: Pathways, barriers and detours in self-directed online music learning », *Journal of Music, Technology & Education*, vol. 5, n° 3, 2013, p. 293-308.
- LEE, Karen V., « A neophyte about online teaching: Almost done », *Qualitative Inquiry*, vol. 14, n° 7, 2008, p. 1180-1186.
- LIEVROUW, Leah A., « The next decade in Internet time: Ways ahead for new media studies », *Information, Communication & Society*, vol. 15, n° 5, 2012, p. 616-638.
- LING, Richard, *Taken for Grantedness: The Embedding of Mobile Communication into Society*, The MIT Press, 2012.

- MABWEAZARA, Hayes Mawindi, « Researching the use of new technologies (ICTs) in Zimbabwean newsrooms: An ethnographic approach », *Qualitative Research*, vol. 10, n° 6, 2010, p. 659-677.
- MARKHAM, Annette N., « Ethic as method, method as ethic: A case for reflexivity in qualitative ICT research », *Journal of Information Ethics*, vol. 15, n° 2, 2006, p. 37-54.
- MARKHAM, Annette N., *Life Online: Researching Real Experience in Virtual Space*, AltaMira Press, 1998.
- MARKHAM, Annette et Elizabeth BUCHANAN, *Ethical Decision-Making and Internet Research: Recommendations from the AoIR Ethics Working Committee (Version 2.0)*, Association of Internet Researchers, 2012.
- MARRES, Noortje, « The redistribution of methods: On intervention in digital social research, broadly conceived », *The Sociological Review*, vol. 60, 2012, p. 139-165.
- MARSHALL, Jon, « Ambiguity, oscillation and disorder: Online ethnography and the making of culture », *Cosmopolitan Civil Societies: An Interdisciplinary Journal*, vol. 2, n° 3, 2010.
- MILLER, Daniel et Heather A. HORST, « The digital and the human: A prospectus for digital anthropology », dans Heather A. HORST et Daniel MILLER (dir.), *Digital Anthropology*, Berg Publishers, 2012, p. 3-38.
- MILLER, Daniel et Don SLATER, *The Internet: An Ethnographic Approach*, Berg Publishers, 2000.
- MONACO, Jeanette, « Memory work, autoethnography and the construction of a fan-ethnography », *Participations*, vol. 7, n° 1, 2010.
- MURTHY, Dhiraj, « Ethnographic Research 2.0: The potentialities of emergent digital technologies for qualitative organizational research », *Journal of Organizational Ethnography*, vol. 2, n° 1, 2013, p. 23-36.
- MURTHY, Dhiraj, « Digital ethnography: An examination of the use of new technologies for social research », *Sociology*, vol. 42, n° 5, 2008, p. 837-855.
- NARDI, Bonnie A., *My Life as a Night Elf Priest: An Anthropological Account of World of Warcraft*, University of Michigan Press, 2010.
- NEMETH, David J. et Rena C. GROPPER, « A cyber-ethnographic foray into GGrandT Internet photo blogs », *Romani Studies*, vol. 18, n° 1, 2008, p. 39-70.
- NICHOLS, Lawrence T. et Jason RINE, « Collective identity narratives: Historical and emergent stories of selfhood in a deindustrializing community », *Studies in Symbolic Interaction*, vol. 39, 2012, p. 121-149.
- ORTON-JOHNSON, Kate, « Knit, purl and upload: New technologies, digital mediations and the experience of leisure », *Leisure Studies*, vol. 33, n° 3, 2014, p. 305-321.
- ORTON-JOHNSON, Kate et Nick PRIOR (dir.), *Digital Sociology: Critical Perspectives*, Palgrave Macmillan, 2013.
- PACCAGNELLA, Luciano, « Getting the seats of your pants dirty: Strategies for ethnographic research on virtual communities », *Journal of Computer-Mediated Communication*, vol. 3, n° 1, 1997.
- PAECHTER, Carrie, « Researching sensitive issues online: Implications of a hybrid insider/outsider position in a retrospective ethnographic study », *Qualitative Research*, vol. 13, n° 1, 2013, p. 71-86.

- PARKS, Malcolm, « What will we study when the Internet disappears? », *Journal of Computer-Mediated Communication*, vol. 14, n° 3, 2009, p. 724-729.
- PARRY, Keith D., « Game of two passions: A football fan's autoethnography », *Qualitative Research Journal*, vol. 12, n° 2, 2012, p. 238-250.
- POSTILL, John et Sarah PINK, « Social media ethnography: The digital researcher in a messy web », *Media International Australia*, vol. 145, n° 1, 2012, p. 123-134.
- REED-DANAHAY, Deborah E. (dir.), *Auto/Ethnography: Rewriting the Self and the Social*, Berg Publishers, 1997.
- ROBINSON, Laura et Jeremy SCHULZ, « New avenues for sociological inquiry: Evolving forms of ethnographic practice », *Sociology*, vol. 43, n° 4, 2009, p. 685-698.
- ROGERS, Richard, *Digital Methods*, The MIT Press, 2013.
- RUPPERT, Evelyn, John LAW et Mike SAVAGE, « Reassembling social science methods: The challenge of digital devices », *Theory, Culture and Society*, vol. 30, n° 4, 2013, p. 22-46.
- SADE-BECK, Liav, « Internet ethnography: Online and offline », *International Journal of Qualitative Methods*, vol. 3, n° 2, 2008, p. 45-51.
- SENF, Theresa M., *Camgirls: Celebrity and Community in the Age of Social Networks*, Peter Lang, 2008.
- SPARKES, Andrew C., « Autoethnography: Self-indulgence or something more? », dans Arthur P. BOCHNER et Carolyn ELLIS (dir.), *Ethnographically Speaking: Autoethnography, Literature, and Aesthetics*, AltaMira Press, 2002, p. 209-232.
- STEINMETZ, Kevin F., « Message received: Virtual ethnography in online message boards », *International Journal of Qualitative Methods*, vol. 11, n° 1, 2012, p. 26-39.
- TRERÉ, Emiliano, « Social movements as information ecologies: Exploring the coevolution of multiple Internet technologies for activism », *International Journal of Communication*, vol. 6, 2012, p. 2359-2377.
- TSCHIDA, Christina M. et Brian R. SEVIER, « Teaching social studies online: An exemplar for examining the broader implications of online methods courses in teacher education », *Journal of Online Learning and Teaching*, vol. 9, n° 4, 2013, p. 500-515.
- TUNCALP, Deniz et Patrick L. LÊ, « (Re)locating boundaries: A systematic review of online ethnography », *Journal of Organizational Ethnography*, vol. 3, n° 1, 2014, p. 59-79.
- UNDERBERG, Natalie M. et Elayne ZORN, *Digital Ethnography: Anthropology, Narrative, and New Media*, University of Texas Press, 2013.
- WILKERSON, J. Michael *et al.*, « Recommendations for Internet-based qualitative health research with hard-to-reach populations », *Qualitative Health Research*, vol. 24, n° 4, 2014, p. 561-74.
- WILKINSON, Catherine et Anthony PATTERSON, « Peppa Piggy in the middle of marketers and mashup makers », dans Stephen BROWN et Sharon PONSONBY-McCABE (dir.), *Brand Mascots: And Other Marketing Animals*, Routledge, 2014, p. 123-140.

CHAPITRE 5

Les bénéfices mutuels de l'ethnographie et de l'analyse du discours en contexte numérique

David Myles

Il existe une tradition ethnographique aussi répandue que diversifiée dans le champ de la communication qui s'est formalisée à partir des années 1970 jusqu'à aujourd'hui. Par exemple, un pan de la littérature a exploré le potentiel analytique et méthodologique de l'ethnographie pour étudier les événements de la vie quotidienne en privilégiant la communication et l'interaction comme principal point d'entrée (De Salins, 1992). Un autre corpus de recherches influentes (Morley et Silverstone, 1991; Silverstone *et al.*, 1991) s'est penché sur les stratégies quotidiennes d'appropriation des médias par les usagers et usagères, donnant entre autres naissance à la théorie de la domestication (Haddon, 2007) et à la sociologie des usages (Jauréguiberry et Proulx, 2011). D'autres études se sont plutôt intéressées au rôle central de la communication dans la constitution d'organisations et de collectifs (Schwartzman, 1993), mobilisant notamment des méthodes d'observation à tradition ethnographique comme le *shadowing* (« filature ») (Meunier et Vásquez, 2008). En effet, l'apport de l'ethnographie au champ de la communication est tel qu'il devient impossible d'en offrir une analyse exhaustive dans le présent chapitre qui, en conséquence, se limite à la transposition des approches ethnographiques en contexte numérique.

Avec l'avènement d'Internet et des technologies numériques, les approches ethnographiques font face à d'importants défis en matière

d'accès, de collecte et d'interprétation de données. Fortement ancrée dans le courant des méthodes virtuelles (Marres, 2012), l'ethnographie du Web (Hine, 2015), que l'on appelle également ethnographie en ligne (Markham, 2005), netnographie (Kozinets, 2009) ou encore ethnographie virtuelle (Hine, 2000), est directement tributaire de ces mutations. L'ethnographie du Web se rapporte à un ensemble de concepts et d'outils principalement qualitatifs mobilisés dans des recherches voulant saisir les phénomènes socioculturels qui sont non seulement liés au développement d'Internet, mais qui souvent émergent sur Internet comme tels (voir Hine, le présent ouvrage). Popularisée par Hine (2000 ; 2015), l'ethnographie du Web englobe une panoplie d'approches, au point qu'il serait plus approprié de parler d'« approches ethnographiques du Web » au pluriel (Coleman, 2010 ; Garcia, 2009). Cela dit, ces approches font généralement appel à au moins deux stratégies de collecte de données, soit l'analyse du discours et l'observation ethnographique (qu'elle soit participante ou non, en ligne ou hors-ligne). Comme l'affirme Hine (traduction libre de 2000, p. 53-54) :

Plutôt que de remplacer l'ethnographie, les approches analytiques du discours utilisées [pour étudier] les textes en ligne peuvent coexister avec les approches ethnographiques [pour étudier] les interactions en ligne. Cette combinaison peut aider à maintenir une ambivalence analytique face à ce que le phénomène à l'étude est vraiment.

Si la combinaison de l'ethnographie et de l'analyse du discours possède en effet un riche potentiel analytique pour l'étude de terrains et d'objets numériques, peu de réflexions ont jusqu'à présent été menées de manière explicite sur les tensions existantes entre ces deux méthodes. Comme le souligne Spencer (traduction libre de 1994, p. 267-268), quoiqu'elles favorisent chacune l'analyse empirique de phénomènes sociaux, l'ethnographie et l'analyse du discours reposent sur des épistémologies bien différentes, la première approche visant à « décrire les contextes sociaux et organisationnels selon la perspective des participants », alors que la seconde privilégie l'analyse « de la structure et du contenu de conversations qui se produisent naturellement ». L'objectif du présent chapitre consiste donc à relever les tensions, mais surtout à explorer les possibilités qu'offre l'arrimage de l'ethnographie et de l'analyse du discours en contexte numérique. Pour ce faire, il se divise en deux sections qui répondent chacune à une question : (1) Que peut apporter l'analyse du discours à l'ethnographie en contexte numérique ?, et (2) Que

peut apporter l'ethnographie à l'analyse du discours en contexte numérique? Nous verrons que la clé d'un arrimage cohérent tant sur les plans conceptuel que méthodologique consiste d'abord à clarifier la visée de chacune des méthodes utilisées, ainsi qu'à expliquer les relations établies entre ces méthodes.

Les apports de l'analyse du discours à l'ethnographie en contexte numérique

Vers la fin des années 1990, des travaux notamment issus de la communication, de l'anthropologie et de la sociologie culturelle ont commencé à s'intéresser aux usages d'Internet au sein de divers groupes sociaux (Robinson et Schulz, 2009). Pour ce faire, ces études ont eu recours à des outils ethnographiques; d'un côté pour comprendre comment les membres de groupes ou de communautés « hors-ligne » s'appropriaient les technologies numériques et intégraient celles-ci dans des pratiques communes préexistantes (Miller et Slater, 2000), d'un autre pour examiner l'émergence même de collectifs « en ligne » et les codes culturels développés par ces derniers en se servant de la communication médiatisée par ordinateur (Baym, 2000; Kendall, 2002). C'est ce dernier corpus de recherches qui a contribué au développement de bonnes pratiques en matière de recherche qualitative menée sur et au moyen d'Internet (Mann et Stewart, 2000), empruntant des méthodes d'inspiration ethnographique comme l'observation participante en ligne et l'entretien qualitatif auprès de personnes informatrices clés (en face à face, par courriel, par messagerie instantanée, etc.), mais proposant aussi des méthodes complémentaires, comme l'analyse discursive, dans le but d'exploiter les traces en ligne.

« À quoi s'intéresse au juste l'ethnographie? » Malgré l'existence d'une grande variété d'approches (Hammersley, 2005), l'ethnographie repose sur une posture générale visant à rendre compte de la complexité de phénomènes socioculturels de manière holistique en favorisant le point de vue de l'acteur comme herméneutique. En d'autres termes, employer une posture ethnographique revient non seulement à chercher à comprendre comment les individus donnent un sens à leur réalité, mais aussi à utiliser leurs catégories et leur vocabulaire comme une clé interprétative pour analyser les phénomènes à l'étude. Comme le souligne Hine (traduction

libre, 2015, p. 1), «l'ethnographie est prisée en tant que méthode pour atteindre le cœur de la signification nous permettant de comprendre, exhaustivement et en profondeur, comment les personnes donnent un sens à leurs vies». Conséquemment, les méthodes privilégiées par l'ethnographe sont celles qui permettent justement de relever les significations que les personnes accordent à leurs actions, d'où l'importance de l'observation participante *in situ* et de l'entretien qualitatif. Ces méthodes permettent en effet d'accéder aux représentations, cognitions et affects des individus (dans une posture plus positiviste) ou de coconstruire des savoirs avec ces derniers (dans une posture plus constructiviste) en effectuant une démarche d'écriture narrative. On attend donc traditionnellement de l'ethnographe une immersion de longue durée au sein d'un terrain, immersion qui vise à développer une compréhension plus fine des codes culturels, normatifs et identitaires d'un collectif ou d'une communauté donnée. En ce sens, les méthodes ethnographiques ont comme premier objectif de «contextualiser» (une culture, des pratiques, des identités, etc.), dans la mesure où l'objet d'analyse «est» le contexte tel qu'il est appréhendé, dans toute sa complexité, ses tensions et ses contradictions.

«Comment situer l'analyse du discours en relation à l'ethnographie?» Malgré leurs origines distinctes, quelques approches combinent l'ethnographie et l'analyse du discours, que ce soit sous l'étiquette de l'ethnographie de la communication, de la microethnographie ou de l'ethnographie critique (Atkinson *et al.*, 2011). De manière générale, les approches ethnographiques utilisent les outils d'analyse discursive pour inscrire concrètement les observations dans un cadre interactionnel spécifique. L'analyse du discours offre la possibilité d'étudier, sur le plan interactionnel et de manière plus systématique, les pratiques langagières d'un groupe ou d'une communauté à l'aide d'outils d'enregistrement (enregistreuse audio, caméra vidéo, etc.). Cela est crucial, dans la mesure où «l'interaction est centrale à l'organisation de la culture et à l'organisation sociale» (traduction libre de Goodwin, 1990, p. 1). Ainsi, la complémentarité de l'analyse du discours réside dans sa propension à rendre compte de la constitution et du reflet d'une culture donnée à travers l'interaction, toujours dans une visée ethnographique de contextualisation. L'analyse du discours constitue l'«un» des outils disponibles (comme l'observation et l'entretien) pour comprendre ce que signifie le fait d'«être» ou de «faire» dans un contexte socioculturel donné. Qui plus est, contrairement à l'entretien qualitatif,

l'analyse du discours est menée sur des données extraites en milieu (plus ou moins, selon les cas) naturels. Elle permet en ce sens d'appréhender les pratiques discursives ou communicationnelles des acteurs comme telles et non selon les conceptions que ces derniers en ont. En effet, si l'entretien qualitatif représente une méthode de choix pour comprendre les affects, les cognitions ou les représentations des individus, se concentrer sur les pratiques langagières permet plutôt de souligner les comportements apparaissant comme significatifs dans l'établissement de routines, ainsi que de mettre en question les conclusions provisoires de l'ethnographe à la suite d'une phase d'immersion (Spencer, 1994).

« Comment cela s'articule-t-il en contexte numérique ? » La combinaison de l'ethnographie et de l'analyse du discours est d'autant plus importante pour la recherche en contexte numérique qu'une grande proportion des pratiques utilisées par les internautes est susceptible de générer des traces pouvant être exploitées dans les recherches (Beaulieu, 2004; Hine, 2015). Grâce aux différentes affordances des plateformes en ligne comme la visibilité et la persistance (Treem et Leonardi, 2012), ces discours se matérialisent sous forme de textes, d'images, de vidéos, de métadonnées, de contenus hybrides, etc. Ainsi, un nombre croissant de recherches d'inspiration ethnographique utilisent aujourd'hui ces traces. Par exemple, Geiger et Ribes (2011, p. 1) proposent l'ethnographie sur traces, qu'ils définissent comme une technique flexible combinant l'exploitation « de la prolifération de documents et de traces documentaires » avec la « richesse de l'observation participante » dans le but de reconstruire à distance les pratiques étudiées sans pour autant nécessiter la coprésence physique entre les individus et l'ethnographe. Ainsi, l'ethnographie sur traces s'opère nécessairement en combinant des outils d'analyse discursive (ou, éventuellement, de tout autre type d'analyse textuelle, comme l'analyse thématique de contenu, l'analyse des interactions, l'analyse conversationnelle, l'analyse sémiotique, etc.) et d'observation participante multisite et délocalisée.

Par multisite et délocalisée, nous entendons que les pratiques des usagers et usagères s'exercent dans une multitude d'espaces tant spatiaux que temporels (en ligne/hors-ligne, sur diverses plateformes parallèles, à distance, en situation de mobilité, etc.), dont les outils de l'ethnographe doivent tenir compte et qu'ils doivent refléter. C'est cette posture holistique chez l'ethnographe du Web, c'est-à-dire sa propension à examiner l'activité

humaine tant dans ses spécificités que dans son ensemble, qui l'amène à formuler des critiques à l'égard des démarches dites ethnographiques reposant exclusivement sur des outils d'observation non participante, sans avoir recours à l'entretien ou à un investissement actif sur le terrain. Comme le soutiennent de nombreux travaux (Beaulieu, 2004; Hine, 2000), les méthodes non obstructives ou non participantes seraient limitées, car celles-ci ne permettraient pas à l'ethnographe de rendre compte des dimensions cachées ou implicites du phénomène ou du groupe à l'étude, d'où l'importance d'entrer en rapport dialogique avec les gens qui participent (Garcia, 2009). Pour Hine (traduction libre de 2008, p. 262), ces stratégies, que l'on désigne communément par l'étiquette péjorative de *lurking* (épier, traquer de loin, se cacher pour observer, etc.), impliqueraient « un manque d'engagement et d'habileté pour développer une compréhension profonde et intérieure que l'ethnographie requiert ». Quoique les pratiques de *lurking* soulèvent d'importants enjeux éthiques sur la recherche en matière de vie privée et de droit à l'autodétermination en contexte numérique (Sveningsson, 2004), nous verrons dans la prochaine section que celles-ci peuvent au contraire constituer une méthode de collecte de données valable et avantageuse en fonction du contexte de la recherche et de ses objectifs.

En résumé, l'ethnographie se distingue principalement de l'analyse du discours dans sa propension à utiliser le point de vue de l'individu ou de la communauté comme principale clé interprétative pour analyser les données d'observation, ainsi que dans son intérêt premier pour le contexte. Dans le cadre d'une ethnographie, procéder à l'analyse discursive d'événements interactionnels permet d'atteindre au moins trois objectifs. D'abord, l'analyse discursive permet de valider les interprétations de l'ethnographe, c'est-à-dire vérifier de manière plus systématique si les conclusions tirées d'observations participantes et immersives concordent avec ce que disent et font les individus. Ensuite, ce type d'analyse permet de souligner les pratiques langagières communes d'un groupe ou d'une communauté. Lors d'une ethnographie, l'analyste s'intéresse généralement – mais pas exclusivement, selon les approches – aux pratiques discursives comme étant le lieu de la performance ou de la représentation d'une culture donnée. Enfin, l'analyse du discours permet de déterminer le vocabulaire utilisé par les personnes, notamment dans le but d'analyser les données d'observation recueillies à la lumière de ces catégories lexicales

et sémantiques. Ainsi, cette première facette de l'arrimage entre l'ethnographie et l'analyse du discours suit une logique « descendante » qui permet d'ancrer les réflexions portant sur un contexte ou sur une culture dans l'accomplissement concret de pratiques discursives.

Les apports de l'ethnographie à l'analyse du discours en contexte numérique

Tout comme l'ethnographie, l'analyse du discours est un terme générique qui fait référence à un ensemble de perspectives et de théories hétéroclites (Hammersley, 2005). Fortement ancrée dans le tournant linguistique en sciences sociales et humaines (Rorty, 1967), l'analyse du discours, dans son acception la plus générale, se distingue de l'ethnographie dans sa propension à vouloir rendre compte de l'ontogenèse des phénomènes sociaux, c'est-à-dire reconnaître les processus qui sont responsables de la constitution, du maintien et des mutations de la réalité sociale. Quoique d'autres théories proposent leur propre vision ontogénique de la réalité sociale et des processus communicationnels assurant son organisation (Luhmann, 1995), l'analyse du discours a comme particularité de faire de la communication sur le plan de l'interaction le principal, voire l'unique point d'accès. L'analyse du discours repose donc sur une analyse systématique (plus ou moins, selon le cas) d'événements interactionnels dans le but de comprendre le contexte social ayant permis l'émergence de ces événements (la dimension représentative), ainsi que la manière dont ces événements participent à l'organisation de la réalité sociale (la dimension constitutive). Elle se distingue notamment de l'analyse thématique de contenu (Bardin, 1977) qui, en calculant et en schématisant les occurrences d'unités sémantiques spécifiques dans un corpus donné, cherche à offrir un portrait se voulant tant qualitativement que quantitativement représentatif du phénomène étudié dans une perspective plus positiviste.

« En quoi consiste l'analyse du discours ? » S'il existe une panoplie d'approches faisant l'objet de nombreuses typologies (Putnam et Fairhurst, 2001; Fairhurst et Putnam, 2014¹; Cooren, 2015), l'analyse du discours

1. Chez Fairhurst et Putnam (2014), ces approches sont la sociolinguistique, l'analyse conversationnelle, la linguistique cognitive, la pragmatique, les actes de langage, l'ethnographie du langage, l'analyse des interactions, la sémiotique, les analyses rhétoriques et littéraires, les analyses critiques ainsi que les analyses postmodernes.

partage avec l'ethnographie la caractéristique consistant à étudier les phénomènes sociaux et humains dans une perspective empirique (à l'exception, peut-être, des approches critiques et postmodernes), c'est-à-dire en partant d'observations de terrain. Si la distinction des caractéristiques de chacune de ces approches va bien au-delà de la portée de ce chapitre, rappelons d'entrée de jeu qu'une démarche analytique du discours peut puiser dans plusieurs traditions complémentaires, tant que son opérationnalisation « est ancrée dans l'action et, plus précisément, dans la communication (comme pratique sociale centrale) » (traduction libre de Cooren *et al.*, 2011, p. 6). Particulièrement pertinente pour l'analyse discursive en contexte numérique, l'analyse conversationnelle (Pomerantz et Fehr, 1997) met en lumière les stratégies utilisées par les acteurs afin d'organiser leurs interactions en examinant certains éléments comme les formules d'ouverture et de fermeture (*openings/closings*), les tours de parole (*turn taking*) ou les relations causales établies entre deux ou plusieurs énoncés (*adjacency pairs*) (Putnam et Fairhurst, 2001). L'analyse conversationnelle repose grandement sur le programme de l'ethnométhodologie (Garfinkel, 1967), qui s'intéresse à la réalisation de stratégies langagières quotidiennes auxquelles les acteurs ont recours pour organiser leur monde (Drew et Heritage, 1992). Ainsi, contrairement à l'ethnographie, l'analyse du discours ne s'intéresse pas *a priori* aux significations que les individus entretiennent (face à leurs comportements, leurs usages, leurs rôles, etc.), mais plutôt aux « conduites signifiantes qui sont produites et saisies par des procédures et des méthodes partagées » (traduction libre de Pomerantz et Fehr, 1997, p. 69). C'est donc dire que l'analyse du discours relate les conséquences que ces significations produisent lorsqu'elles sont mises en acte dans des routines (par des paroles, des gestes, etc.). À ce titre, l'analyse conversationnelle rejoint les approches pragmatiques du discours et la théorie sur les actes de langage (Searle, 1969), qui s'intéressent également à la structure interactionnelle, mais plus encore à la performativité des discours, c'est-à-dire leur propension à produire des effets, à faire une différence significative dans une interaction.

« Comment s'opérationnalise l'analyse du discours ? » Comme les pratiques discursives des acteurs participant à l'organisation du quotidien sont souvent déployées de manière inconsciente (c'est-à-dire qu'elles s'opèrent à leur insu), les analystes du discours ont tendance à privilégier des méthodes de collecte de données en milieux dits naturels (par des

enregistrements audiovisuels) et à délaissier les méthodes visant à (co)produire des données ou risquant de modifier les comportements des acteurs (comme l'observation participante et l'entretien) (Spencer, 1994; Maynard, 2006). En ce sens, les traces laissées en ligne par les activités d'internautes constituent l'occasion d'analyser le discours et de comprendre comment s'organisent les interactions et, par extension, comment se constitue la réalité sociale, puisque Internet permet une collecte non obstructive et passablement rapide de données qualitatives potentiellement riches. Cela dit, l'observation participante et l'entretien demeurent pertinents pour l'analyste du discours (Hammersley, 2005), dans la mesure où, même si elles ne servent pas à produire les données principalement analysées, ces méthodes permettent de « contextualiser » l'événement interactionnel à l'étude. Ainsi, dans une logique « ascendante », l'entretien mené auprès d'acteurs dont les traces ont été recueillies en ligne ainsi que l'observation participante *a posteriori* auprès de ces derniers amènent un plus grand contrôle analytique. Par contrôle analytique, nous faisons référence à la capacité de déceler les informations contextuelles qui vont au-delà du cadre immédiat de l'énoncé ou de l'interaction, mais qui sont nécessaires à son interprétation (Maynard, 2006). Si l'entretien et l'observation participante immersive sont déterminantes pour la démarche ethnographique, elles constituent des stratégies complémentaires pour l'analyse du discours (tout comme l'observation non participante, la recherche documentaire en ligne, la revue de littérature scientifique, etc.), permettant surtout de contextualiser les données.

« Comment l'analyse du discours s'applique-t-elle en contexte numérique ? » Comparativement à l'ethnographie, il n'existe pas à ce jour de corpus de recherches robuste se penchant spécifiquement sur les possibilités de l'analyse du discours en contexte numérique, malgré l'émergence dans la francophonie de travaux hétérogènes utilisant l'analyse conversationnelle (Marcochia, 2016) ou discursive (Paveau, 2017) sur des terrains en ligne. L'un des principaux défis de l'analyse du discours en ligne est sans doute d'éviter la reproduction du clivage traditionnellement établi entre les dimensions soi-disant symboliques et matérielles du discours (Putnam, 2015²). L'une des pistes de solution se trouve notamment dans

2. Dans cet article, Putnam (2015) invite à rejeter tout déterminisme matériel (le fait de surestimer l'importance des objets physiques ou numériques) et tout relativisme constructiviste (le fait de surestimer l'importance des processus symboliques) dans notre

la théorie de l'acteur-réseau (Latour, 2005) et dans les approches constitutives de la communication (dites CCO, Brummans *et al.*, 2014). Selon la théorie de l'acteur-réseau, les processus d'organisation (ou *organizing*) de la réalité sociale s'opèrent par la distribution d'agentivité entre divers acteurs à « ontologies variables » (Law, 1994). Par agentivité, nous entendons la capacité de faire une différence significative en interaction (Cooren, 2013). Dans cette perspective, les processus d'*organizing* s'opèrent par le développement de pratiques discursives (et, par extension, par la communication) qui mettent en scène des acteurs tant humains que non humains. Cela ne veut pas dire que l'agentivité humaine n'est pas importante, mais que celle-ci a peut-être été jusqu'à présent surestimée au détriment d'entités non humaines (Cooren, 2006), une croyance qui diverge des perspectives ethnographiques traditionnelles centrées sur l'humain (à l'exception, sans doute, de l'anthropologie de la technique). En contexte numérique, les acteurs non humains peuvent entre autres prendre la forme d'affordances, d'algorithmes et d'architecture numérique ou encore, de figures rhétoriques, de métaphores et d'idéologies qui, par ailleurs, finissent toujours par être matérialisées. Ainsi, en tenant compte de l'agentivité humaine et non humaine, ces perspectives ont le potentiel de souligner de manière explicite la nature intrinsèquement sociomatérielle (Orlikowski, 2007) des processus d'*organizing* en contexte numérique en réfutant l'ontologie de la séparation établie entre socialité et matérialité (Barad, 2003).

Les théories portant sur les affordances de la matérialité numérique (Bucher et Helmond, 2017), c'est-à-dire ce qu'offrent les technologies ou médias numériques en matière de possibilités et ce qu'ils imposent en matière de contraintes pour l'action, semblent particulièrement riches pour comprendre les processus de matérialisation discursive, notamment lorsqu'elles sont appréhendées dans une perspective d'*organizing* (Leonardi et Vaast, 2017). Ces théories sont cohérentes avec celle de l'acteur-réseau, dans la mesure où elles reposent également sur une ontologie relationnelle stipulant que l'agentivité n'est pas située chez les acteurs eux-mêmes, mais

compréhension de la réalité sociale et de son organisation. Pour ce faire, elle examine des modèles alternatifs (notamment issus des études organisationnelles), qui conceptualisent à leur manière le « discours » considéré comme étant à la fois constitué de dimensions symboliques et matérielles.

plutôt distribuée par l'entremise de ces derniers (Faraj et Azad, 2012). Ainsi, toute approche analytique du discours en contexte numérique aurait avantage à considérer la façon dont les affordances des plateformes numériques sont mises en acte (ou « énoncées ») par l'accomplissement social de pratiques langagières (Jarzabkowski et Pinch, 2013). Il importe donc d'examiner le rôle des plateformes et de leurs affordances dans la matérialisation différenciée des discours et les effets que ces éléments sociomatériels génèrent pour l'organisation quotidienne de la réalité sociale. En ce sens, l'attribution de l'agentivité, c'est-à-dire le fait de reconnaître à quelqu'un ou à quelque chose la capacité de faire une différence productive en interaction, constitue une « question ouverte » (Cooren *et al.*, 2011), question à laquelle seul l'arrimage entre l'analyse du discours et l'observation ethnographique, qu'elle soit menée de manière non obstructive ou participante *a posteriori*, peut répondre de façon éloquente et mutuellement bénéfique.

* * *

Ce chapitre offre une synthèse réflexive des tensions et des possibilités que recèle l'association de l'ethnographie et de l'analyse du discours en contexte numérique. Étant donné les contraintes inhérentes au format du présent chapitre, cette synthèse a été effectuée en simplifiant (parfois à outrance) les spécificités et la multitude des approches ethnographiques et analytiques du discours. En effet, des critiques pourraient être légitimement formulées quant à la possibilité même d'appréhender ces deux approches en des termes si rudimentaires et contrastants. Cela dit, l'objectif de ce chapitre n'était pas d'offrir un survol exhaustif, mais plutôt d'insister sur la complémentarité des deux approches, que ce soit dans une logique descendante (où l'analyse du discours permet d'ancrer plus systématiquement et de valider les observations menées par l'ethnographe) ou dans une logique ascendante (où l'entretien et l'observation ethnographiques permettent de comprendre le contexte d'émergence d'un événement interactionnel, ainsi que d'acquérir un meilleur contrôle analytique). Ultimement, les méthodes privilégiées par les deux approches possèdent chacune leurs avantages et leurs limites. En ce sens, la sélection d'une méthode devrait d'abord reposer sur la cohérence de celle-ci avec les objectifs et le cadrage théorique de la recherche qui la mobilise et non sur

la croyance voulant que certaines méthodes soient naturellement supérieures ou inférieures à d'autres. En misant sur la triangulation des méthodes et, surtout, sur l'explication des liens entre celles-ci (c'est-à-dire sur ce qu'une méthode permet d'accomplir en relation à une autre), l'arrimage de l'ethnographie et de l'analyse du discours en contexte numérique a tout le potentiel pour offrir un éclairage riche afin de saisir, dans une perspective résolument compréhensive, les mutations liées au développement des technologies numériques et à leurs usages.

Références

- ATKINSON, Dwight, Hanako OKADA et Steven TALMY, « Ethnography and discourse analysis », dans Ken HYLAND et Brian PALTRIDGE (dir.), *Continuum Companion to Discourse Analysis*, Continuum Books, 2011, p. 85-100.
- BARAD, Karen, « Posthumanist performativity: Toward an understanding of how matter comes to matter », *Signs: Journal of Women in Culture and Society*, vol. 28, n° 3, 2003, p. 801-831.
- BARDIN, Laurence, *L'analyse de contenu*, Presses universitaires de France, 1977.
- BAYM, Nancy K., *Tune In, Log On: Soaps, Fandom, and Online Community*, SAGE Publications, 2000.
- BEAULIEU, Anne, « Mediating ethnography: Objectivity and the making of ethnographies of the Internet », *Social Epistemology*, vol. 18, nos 2-3, 2004, p. 139-163.
- BRUMMANS, Boris H. J. M. et al., « Approaches to research on the communicative constitution of organizations », dans Linda L. PUTNAM et Dennis K. MUMBY (dir.), *The SAGE Handbook of Organizational Communication: Advances in Theory, Research, and Methods*, SAGE Publications, 2014, p. 173-194.
- BUCHER, Taina et Anne HELMOND, « The affordances of social media platforms », dans Jean BURGESS, Alice MARWICK et Thomas POELL (dir.), *The SAGE Handbook of Social Media*, SAGE Publications, 2017, p. 233-253.
- COLEMAN, E. Gabriella, « Ethnographic approaches to digital media », *Annual Review of Anthropology*, vol. 39, 2010, p. 487-505.
- COOREN, François, *Organizational Discourse: Communication and Constitution*, John Wiley & Sons, 2015.
- COOREN, François, *Manières de faire parler: interaction et ventriloquie*, Le Bord de l'eau, 2013.
- COOREN, François, « The organizational world as a plenum of agencies », dans François COOREN, James R. TAYLOR et Elizabeth J. VAN EVERY, *Communication as Organizing: Practical Approaches to Research into the Dynamic of Text and Conversation*, Lawrence Erlbaum Associates, 2006, p. 81-100.
- COOREN, François et al., « Communication, organizing and organization: An overview and introduction to the special issue », *Organization Studies*, vol. 32, n° 9, 2011, p. 1149-1170.

- DE SALINS, Geneviève-Dominique, *Une introduction à l'ethnographie de la communication*, Didier, 1992.
- DREW, Paul et John HERITAGE, *Talk at Work: Interaction in Institutional Settings*, Cambridge University Press, 1992.
- FAIRHURST, Gail T. et Linda L. PUTNAM, « Organizational discourse analysis », dans Linda L. PUTNAM et Dennis K. MUMBY (dir.), *The SAGE Handbook of Organizational Communication: Advances in Theory, Research, and Methods*, SAGE Publications, 2014, p. 271-296.
- FARAJ, Samer et Bijan AZAD, « The materiality of technology: An affordance perspective », dans Paul LEONARDI, Bonnie NARDI et Jannis KALLINIKOS (dir.), *Materiality and Organizing: Social Interaction in a Technological World*, Oxford University Press, 2012, p. 237-258.
- GARCIA, Angela C. *et al.*, « Ethnographic approaches to the Internet and computer-mediated communication », *Journal of Contemporary Ethnography*, vol. 38, n° 1, 2009, p. 52-84.
- GARFINKEL, Harold, *Studies in Ethnomethodology*, Prentice-Hall, 1967.
- GEIGER, R. Stuart et David RIBES, « Trace ethnography: Following coordination through documentary practices », dans HICSS '11 *Proceedings of the 2011 44th Hawaii International Conference on System Sciences*, 2011, p. 1-10.
- GOODWIN, Marjorie H., *He-Said-She-Said: Talk as Social Organization Among Black Children*, Indiana University Press, 1990.
- HADDON, Leslie, « Roger Silverstone's legacies: Domestication », *New Media & Society*, vol. 9, n° 1, 2007, p. 25-32.
- HAMMERSLEY, Martyn, « Ethnography and discourse analysis: Incompatible or complementary? », *Polifonia*, vol. 10, n° 10, 2005, p. 1-20.
- HINE, Christine, *Ethnography for the Internet: Embedded, Embodied and Everyday*, Bloomsbury Publishing, 2015.
- HINE, Christine, « Virtual ethnography: Modes, varieties, affordances », dans Nigel G. FIELDING, Raymond M. LEE et Grant BLANK (dir.), *The SAGE Handbook of Online Research Methods*, SAGE Publications, 2008, p. 257-270.
- HINE, Christine, *Virtual Ethnography*, SAGE Publications, 2000.
- JARZABKOWSKI, Paula et Trevor PINCH, « Sociomateriality is "the New Black": Accomplishing repurposing, reinscripting and repairing in context », *M@n@gement*, vol. 16, n° 5, 2013, p. 579-592.
- JAUÉGUIBERRY, Francis et Serge PROULX, *Usages et enjeux des technologies de communication*, Érès, 2011.
- KENDALL, Lori, *Hanging Out in the Virtual Pub: Masculinities and Relationships Online*, University of California Press, 2002.
- KOZINETS, Robert V., *Netnography: Doing Ethnographic Research Online*, SAGE Publications, 2009.
- LATOUR, Bruno, *Reassembling the Social: An Introduction to Actor-Network-Theory*, Oxford University Press, 2005.
- LAW, John, *Organizing Modernity: Social Ordering and Social Theory*, Blackwell, 1994.

- LEONARDI, Paul et Emmanuelle VAAST, «Social media and their affordances for organizing: A review and agenda for research», *Academy of Management Annals*, vol. 11, n° 1, 2017, p. 150-188.
- LUHMANN, Niklas, *Social Systems*, Stanford University Press, 1995.
- MANN, Chris et Fiona STEWART, *Internet Communication and Qualitative Research: A Handbook for Researching Online*, SAGE Publications, 2000.
- MARCOCCIA, Michel, *Analyser la communication numérique écrite*, Armand Colin, Paris, 2016.
- MARKHAM, Annette N., «The methods, politics, and ethics of representation in online ethnography», dans Norman K. DENZIN et Yvonna S. LINCOLN (dir.), *The SAGE Handbook of Qualitative Research*, SAGE Publications, 2005.
- MARRES, Noortje, «The redistribution of methods: On intervention in digital social research, broadly conceived», *The Sociological Review*, vol. 60, 2012, p. 139-165.
- MAYNARD, Douglas, «Ethnography and conversation analysis: What is the context of an utterance?», dans Sharlene Nagy HESSE-BIBER et Patricia LEAVY (dir.), *Emergent Methods in Social Research*, SAGE Publications, 2006, p. 55-94.
- MEUNIER, Dominique et Consuelo VÁSQUEZ, «On shadowing the hybrid character of actions: A communicational approach», *Communication Methods and Measures*, vol. 2, n° 3, 2008, p. 167-192.
- MILLER, Daniel et Don SLATER, *The Internet: An Ethnographic Approach*, Berg Publishers, 2000.
- MORLEY, David et Roger SILVERSTONE, «Communication and context: Ethnographic perspectives on the media audience», dans Klaus BRUHN JENSEN (dir.), *A Handbook of Qualitative Methodologies for Mass Communication Research*, Routledge, 1991, p. 149-162.
- ORLIKOWSKI, Wanda J., «Sociomaterial practices: Exploring technology at work», *Organization Studies*, vol. 29, n° 9, 2007, p. 1435-1448.
- PAVEAU, Marie-Anne, *L'analyse du discours numérique. Dictionnaire des formes et des pratiques*, Hermann, 2017.
- POMERANTZ, Anita et Barbara J. FEHR, «Conversation analysis: An approach to the study of social action as sense making practices», dans Teun A. VAN DIJK (dir.), *Discourse as Social Interaction*, SAGE Publications, 1997, p. 64-91.
- PUTNAM, Linda L., «Unpacking the dialectic: Alternative views on the discourse-materiality relationship», *Journal of Management Studies*, vol. 52, n° 5, 2015, p. 706-716.
- PUTNAM, Linda L. et Gail T. FAIRHURST, «Discourse analysis in organizations», dans Fredric M. JABLON et Linda L. PUTNAM (dir.), *The New Handbook of Organizational Communication: Advances in Theory, Research and Methods*, SAGE Publications, 2001, p. 78-136.
- ROBINSON, Laura et Jeremy SCHULZ, «New avenues for sociological inquiry: Evolving forms of ethnographic practice», *Sociology*, vol. 43, n° 4, 2009, p. 685-698.
- RORTY, Richard (dir.), *The Linguistic Turn: Essays in Philosophical Method*, University of Chicago Press, 1967.

- SCHWARTZMAN, Helen B., *Ethnography in Organizations*, vol. 27, SAGE Publications, 1993.
- SEARLE, John R., *Speech Acts: An Essay in the Philosophy of Language*, Cambridge University Press, 1969.
- SILVERSTONE, Roger, Eric HIRSCH et David MORLEY, « Listening to a long conversation: An ethnographic approach to the study of information and communication technologies in the home », *Cultural Studies*, vol. 5, n° 2, 1991, p. 204-227.
- SPENCER, J. William, « Mutual relevance of ethnography and discourse », *Journal of Contemporary Ethnography*, vol. 23, n° 3, 1994, p. 267-279.
- SVENINGSSON, Malin, « Ethics in Internet ethnography », dans Elizabeth A. BUCHANAN (dir.), *Readings in Virtual Research Ethics: Issues and Controversies*, Information Science Publishing, 2004, p. 45-61.
- TREEM, Jeffrey W. et Paul M. LEONARDI, « Social media use in organizations: Exploring the affordances of visibility, editability, persistence, and association », *Annals of the International Communication Association*, vol. 36, n° 1, 2013, p. 143-189.

CHAPITRE 6

Les terrains de recherche en ligne et hors-ligne : proposition pour une triangulation des méthodes

Sandrine Roginsky

À partir d'une recherche longitudinale menée depuis 2010 sur l'usage des réseaux sociaux numériques en politique, on interroge ici les apports d'une triangulation des méthodes de collecte et d'analyse de données. Cette recherche s'intéresse aux motivations, aux perceptions et aux expériences d'usage des dispositifs¹ sociaux numériques Twitter et Facebook par des personnels politiques, en l'occurrence des membres du Parlement européen et leurs équipes, mais également, plus ponctuellement, des candidates et des candidats à l'élection européenne. Il s'agit notamment de déterminer si, et comment, les représentations et les usages des dispositifs sociaux numériques participent à la reconfiguration des pratiques professionnelles du monde politique. L'usage de dispositifs sociaux numériques est un objet d'étude complexe qui nécessite de ne pas se concentrer uniquement sur ce que nous donnent à voir ces dispositifs pour s'intéresser également à ce qu'ils ne montrent pas, ce qui implique de passer de l'autre côté de l'écran pour accéder aux acteurs et à leurs contextes de production.

1. Dispositif est ici entendu comme artefact sociotechnique qui comprend tout à la fois « une dimension technique, dans le sens où il fournit des formats d'expression; pragmatique, dans le sens où il comprend les interactions qu'il autorise; et enfin symbolique dans la mesure où il comprend un ensemble de significations dégageant différentes portées axiologiques » (Mabi, 2016).

La triangulation des méthodes se révèle utile pour dépasser les limites posées par la seule collecte de données numériques en associant différentes méthodes qualitatives et en combinant ainsi différentes sources de données.

Avec des exemples précis, nous montrerons en quoi le croisement et la combinaison de méthodes, et non leur simple juxtaposition, enrichissent l'analyse. La recherche sur laquelle s'appuie cette réflexion méthodologique justifie le choix de la triangulation. La description des différentes méthodes triangulées et le détail des différents types de données qu'elles ont permis de collecter permettent d'envisager leurs modalités d'analyse. Enfin, le bricolage peut s'avérer nécessaire dès lors que la triangulation amène à combiner différentes méthodes puisque la nature des matériaux collectés exige de travailler des assemblages complexes et mouvants ; ce qui nous amène, finalement, à interroger le statut des données dans une telle démarche de recherche, mais aussi celui du contexte d'action, changeant et difficile à saisir.

Les raisons de la triangulation des méthodes

Les dispositifs socionumériques, comme Facebook et Twitter, sont l'objet d'une « profusion de discours provenant de spécialistes du marketing et de la communication sur Internet, de journalistes et de personnalités politiques » (Coutant et Stenger, 2012, p. 76). Les universitaires ne sont pas en reste. Dans le champ de la communication politique, les travaux prennent généralement pour objet d'étude les productions discursives publiées sur les dispositifs, parvenant parfois à collecter des milliers, voire des millions de messages produits sur la plateforme Twitter. Dans de telles approches, les individus comme les contextes – qu'ils soient sociaux, techniques, institutionnels, politiques, etc. – disparaissent derrière les messages. Les dispositifs également peuvent être éclipsés. Il s'agit ici d'adopter une démarche inverse en nous intéressant à la perspective des acteurs sociaux qui utilisent les dispositifs. On se focalise davantage sur la pratique entendue comme « les comportements, les attitudes et les représentations des individus qui se rapportent directement ou indirectement à l'outil » (Jouët, 1993, p. 371, citée dans Millerand, 1998). Une telle approche exige que l'on se déplace entre les contextes de production et d'usage, autrement dit des deux côtés de l'écran, du « en ligne » au « hors-

ligne » (et réciproquement), puisque les activités se déclinent sur ces deux terrains².

De fait, des jeux de données différents justifient le choix de trianguler les méthodes de collecte. La triangulation donne la possibilité de cumuler et d'articuler différentes méthodes de recueil de données dans le but d'expliquer de manière plus exhaustive la richesse et la complexité des comportements humains (Cohen et Manion, 1986). La triangulation ne doit pas être confondue avec les méthodes mixtes : il ne s'agit pas pour nous de trianguler les méthodes qualitatives et quantitatives, mais de rester dans le paradigme qualitatif, celui de l'enquête de terrain. Pour Beaud et Weber, c'est « la force de l'enquête de terrain que de pouvoir multiplier les points de vue, de faire converger des résultats d'enquête sur un même lieu ou à propos d'un même objet, de pouvoir recouper les informations obtenues » (2010, p. 217).

La triangulation et la mobilité entre les deux côtés de l'écran

Le type de triangulation mobilisé par notre recherche s'ancre dans le paradigme ethnographique. L'objectif de l'ethnographie est de comprendre comment les gens mettent en forme et donnent du sens au monde social dans lequel ils évoluent. L'ethnographie désigne les « méthodes de production des données propres à l'enquête de terrain » (De Sardan, 2008, p. 37). Le terrain est ici appréhendé dans un sens pluriel, intégrant à la fois les terrains « en ligne » et « hors-ligne », qui sont d'ailleurs connectés et imbriqués (Leander et McKim, 2003). L'approche peut ainsi s'apparenter à celle de l'ethnographie multisite qui « ne constitue pas une délocalisation de l'objet de recherche, mais sa localisation plurielle dans un contexte d'intense circulation des acteurs et des constructions socioculturelles » (Dumont, 2014, p. 195). Pour chaque terrain, des protocoles spécifiques de collecte de données ont été mis en place, lesquels ne sont pas sans partager, néanmoins, des caractéristiques communes. Par ailleurs, la recherche étant de nature longitudinale, ceux-ci ont été déployés sur différentes séquences tempo-

2. Le terrain renvoie à « une longue tradition anthropologique » (Beaud et Weber, 2010, p. 5), il est « une forme particulière d'enquête en sciences sociales » (De Sardan, 2008, p. 20). Il est entendu comme « le lieu central de la production des données » (*ibid.*). Il est ainsi « lieu d'investigation, lieu de rencontre [...] défini par une question de recherche, délimité et décelé par les contours, pas nécessairement définitifs, de cette dernière » (Steck, 2012, p. 81).

relles. Ainsi se sont succédé une participation observante sur les terrains « hors-ligne » et « en ligne », une observation sur le terrain « en ligne » avec des entretiens semi-directifs et une extraction de données produites sur le terrain « en ligne », comme le résume le tableau 6.1 ci-dessous.

TABLEAU 6.1

Planification de la triangulation des méthodes

	Participation observante sur le terrain « hors ligne »	Participation observante sur le terrain « en ligne »	Observation sur le terrain « hors ligne »	Entretiens semi-directifs	Extraction de données produites sur le terrain « en ligne »
Août 09 - Déc. 10					
Janv. 11 - Déc. 11					
Janv. 12 - Mai 12					
Juin 12 - Sept. 13					
Oct. 13 - auj.					

Sur les deux terrains, « en ligne » (sur l'écran) et « hors-ligne » (derrière l'écran), nous avons collecté des données au moyen de l'observation, participante ou non. L'observation participante est « le cœur même du terrain ethnographique », selon De Sardan (2001). Elle est menée « au plus près des situations dans lesquelles interviennent les pratiques » (Fournier, 2010, p. 37). Quand elle donne la priorité à la participation et ainsi à la production de données expérientielles, on peut alors parler d'« observation participante » (Soulé, 2007) pour souligner l'intensité de l'engagement et de la participation. L'observation non participante, quant à elle, porte sur « des traces laissées par les pratiques » (Fournier, 2010, p. 37).

L'observation participante

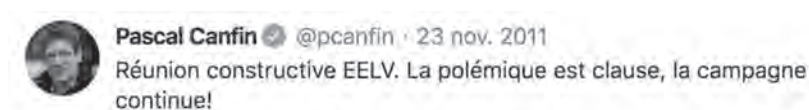
Dans le premier temps de la recherche, et sur une période de trois années, l'observation participante menée sur le terrain « hors-ligne » a été une participation observante. Il s'est alors agi d'utiliser un rôle existant (celui d'atta-

chée parlementaire) pour engager la recherche dans un environnement devenu familier. Cette période de participation observante a rendu possible la compréhension fine du contexte d'activités dans lequel viennent s'insérer les usages des dispositifs ainsi que les postures et les attentes des personnes mobilisées pour les utiliser. La posture s'est également déclinée sur le terrain « en ligne », avec l'écriture de messages sur les dispositifs au nom du député dans le cadre de la fonction d'attachée parlementaire. Elle a permis de saisir plus précisément la multiple nature des contraintes de l'exercice. Elle a également permis d'appréhender un usage des dispositifs qui est généralement peu documenté, voire oublié dans la littérature, celui de l'écoute et de la veille informationnelle en milieu politique (Roginsky, 2015).

Ces données d'observation ont permis de comprendre le contexte de production dans sa dimension multiple et complexe (temporelle, médiatique, politique, institutionnelle, professionnelle, etc.), ce qui a amené à constater le risque de sa disparition si l'on s'en tient uniquement aux données produites sur l'écran pour mener la recherche. À titre d'illustration, ce message publié sur Twitter (figure 6.1) ne révèle pas les coulisses de sa production, et notamment les interactions qu'il suscite dans l'équipe du député ou bien encore avec des journalistes. Qui plus est, sorti de son contexte temporel et médiatique, il est difficile de saisir l'événement politique auquel il se réfère.

FIGURE 6.1

Copie d'écran d'un « tweet »



Cet exemple est intéressant, car il illustre l'entremêlement entre terrain « en ligne » et terrain « hors-ligne » sans pour autant le donner à voir. Autrement dit, les données récoltées uniquement en ligne ne disent rien non seulement des contextes de réception, mais également des contextes de production. Elles ne se substituent pas non plus à une analyse discursive qui n'a pas nécessairement besoin de données supplémentaires pour tenter une interprétation de la phrase en question. Néanmoins, dans ce dernier cas, comment trouver le sens de la phrase sortie de sa séquence temporelle ? Ainsi, à quelle réunion le *tweet* fait-il référence ? À quelle polémique est-il

fait allusion? On saisit par ces questions la complexité de l'analyse de ce type de données qui, bien que déclinées « en ligne », se rapportent à des événements qui ont lieu « hors-ligne ». La question de l'articulation entre le « en ligne » et le « hors-ligne » demeure donc centrale.

De l'observation participante à l'observation

La participation observante s'est terminée en mai 2012, avec la fin du contrat d'attachée parlementaire au Parlement européen. L'observation a alors changé de nature, en abandonnant la dimension participante. « En ligne », elle s'est focalisée sur « l'observation des échanges réguliers et des battements successifs d'interactions, construction des postures et des figures de communication sur les dispositifs » (Roginsky et Huys, 2015, par. 6) « Hors-ligne », les relations construites dans la première période de la participation observante ont permis d'accéder, ponctuellement, à des activités organisées dans l'enceinte du Parlement européen, par exemple des formations proposées par des représentants de Facebook et

FIGURE 6.2

Copies d'écran de « tweets » sur la formation dans l'enceinte du Parlement européen



de Twitter aux personnels politiques. Celles-ci se sont avérées particulièrement intéressantes en cela qu'elles ont permis de saisir au plus près les discours que les firmes leur adressent et les réactions qu'elles suscitent chez ces derniers.

Cette observation des moments de formation dans l'enceinte du Parlement européen a été complétée par l'observation de leur mise en visibilité sur les dispositifs :

Là encore, on voit l'articulation entre les données récoltées sur les terrains « en ligne » et « hors-ligne » : il n'aurait pas été possible d'identifier ces messages sur Twitter, voire d'en comprendre le sens, par la seule observation « en ligne », s'il n'y avait eu au départ l'observation « hors-ligne » des séances de formation.

De l'observation à l'entretien

La technique des entretiens permet de compléter et d'enrichir les données collectées par l'observation. L'entretien reste un moyen privilégié « pour produire des données discursives donnant accès aux représentations populaires, autochtones, indigènes, locales » (De Sardan, 2008, p. 54). En permettant de recueillir des informations sur les logiques des individus et sur leurs perceptions, l'entretien joue différentes fonctions dans le cadre de cette recherche :

- Pendant la période d'observation participante, il a facilité le travail de réflexivité, le questionnement et la mise à distance de l'expérience personnelle.
- Depuis la fin de la période de l'observation participante, il permet de compléter l'observation qui n'est désormais plus de nature participante en recueillant la perspective des individus dans un environnement continuellement changeant.

En résonance avec les rôles qu'il joue, l'entretien présente plusieurs défis :

- Pendant la période d'observation participante, il n'était pas sans poser de difficulté de positionnement puisque la chercheuse qui interrogeait était aussi une collègue, voire une concurrente. Cependant, cette

période lui a également permis d'établir au préalable une relation de confiance indispensable avec les personnes qu'elle allait interviewer.

- Depuis la fin de la période d'observation participante, le défi est d'avoir accès aux acteurs pour les interroger. Le monde politique est un « milieu difficile » à étudier, difficile d'accès (Laurens, 2007). Néanmoins, l'expérience passée, les réseaux constitués grâce à celle-ci sont un atout considérable dans la mise en relation avec les individus ciblés.

D'autres défis, plus génériques, se posent dès lors qu'on mobilise l'entretien pour collecter des données. Il peut en effet être difficile de faire parler des responsables politiques « habitués à donner à leur parole un caractère public » et qui « produisent un discours de représentation d'eux-mêmes qui prend la forme du didactisme » (Chamboredon *et al.*, 1994, p. 128). C'est la raison pour laquelle le choix a été fait d'interroger prioritairement les personnes qui travaillent auprès d'eux, d'autant plus que ce sont ces personnes qui gèrent le quotidien des activités professionnelles – notamment celles qui touchent la communication.

La nécessaire prise en compte des données en ligne

Se limiter aux approches ethnographiques sur le terrain « hors-ligne » pour étudier les usages des dispositifs, c'est passer paradoxalement à côté des dispositifs. Or, comme l'observe Van Dijck (2013), un dispositif socio-numérique ne se contente pas de faciliter les activités, il les dessine également. Partir d'une analyse descriptive des dispositifs, qui s'inspire de l'approche sémiologique, est à ce titre une manière de prendre en compte la place du dispositif dans les interactions comme dans la production de contenu. On voit ci-dessous différentes manières de jouer avec les affordances³ du dispositif par des députées européennes, alors candidates à leur réélection.

L'approche sémiologique cherche à « décrire et à qualifier la mise en signes opérée pour produire, éditer, lire, écrire et documenter des contenus sur le Web » (Bonaccorsi, 2013, p. 140). Elle s'intéresse ainsi à l'environnement numérique, autrement dit à ce que la matérialité des dispositifs donne

3. Les affordances sont des « propriétés de l'environnement et des objets qui s'y trouvent » (Kavanagh *et al.*, 2016), qui définissent les possibilités d'interactions avec ledit environnement.

FIGURE 6.3

Copies d'écran de présentations biographiques sur Twitter

à voir. En cela, elle nous permet d'éclairer les données récoltées au cours de l'observation « en ligne ». Toutefois, se limiter à l'observation « en ligne », c'est aussi passer à côté de la production de contenu que les individus publient sur ces supports. L'observation est en effet terriblement contrainte, le flux géré par les algorithmes ne permettant pas de « tout » voir. Voilà la raison pour laquelle l'extraction automatique de données produites sur les dispositifs peut s'avérer utile. Cet outil de recueil des données s'est ainsi ajouté aux précédents, en se limitant à ce stade au dispositif Twitter. Nous avons travaillé sur des corpus⁴ de *tweets* recueillis par fichier Excel depuis les comptes d'acteurs politiques et institutionnels repérés au préalable grâce à l'observation « hors-ligne » et « en ligne ». Les *tweets* ont été aspirés par un collègue informaticien qui a développé son propre logiciel pour solliciter l'API de Twitter. Notons que cette méthode de collecte est conditionnée par les entreprises propriétaires des dispositifs, puisque ce sont

4. Le corpus est un « ensemble de matériaux empiriques constitués selon une logique précise », guidée par la problématique et les questions de recherche (Mercklé, 2010, p. 29).

elles qui décident ou non d'ouvrir leur application de manière qu'il soit possible, ou non, de collecter des données qui y sont produites. Si lors d'une première extraction entre octobre 2013 et juin 2014, il a été ainsi possible de récolter tous les *tweets* depuis la création des comptes sélectionnés, l'extraction a été moins étendue dans la durée par la suite, l'API restreignant les capacités de remonter dans le temps.

Des données hétérogènes

La triangulation des méthodes nous a menée à récolter différents types de données : (1) les données produites au moyen de l'observation (participante ou non) sur le terrain « hors-ligne » ; (2) les retranscriptions des entretiens semi-directifs suscités et recueillis par nos soins ; (3) les données produites grâce à l'observation (participante ou non) sur le terrain « en ligne » ; (4) les données textuelles extraites automatiquement depuis des comptes sélectionnés sur Twitter ; auxquelles nous pouvons ajouter (5) des données de nature documentaire (textes législatifs, documents institutionnels du Parlement européen, etc.) et journalistique (articles de journaux). En cela, nous nous situons complètement dans le paradigme de la recherche qualitative qui utilise souvent une grande variété de données (Glaser et Strauss, 1967). Ces données hétéroclites sont alors intégrées dans le corpus empiriquement constitué.

La mise en données

Nous entendons par données « la transformation en traces objectivées de “morceaux de réel”, de fragments du réel de référence, tels qu'ils ont été sollicités, sélectionnés et perçus par le chercheur » (De Sardan, 2008, p. 50). Les méthodes de collecte des données ont par conséquent une incidence sur le type de données qu'elles font émerger. L'observation participante permet tout à la fois d'accéder au contexte de production des données et de produire des données sur les interactions en contexte. Les données produites par les entretiens semi-directifs permettent quant à elles de rendre compte du point de vue des acteurs. Dès lors que les entretiens sont retranscrits, ils ont par ailleurs « la particularité de “fixer” ou de “geler” sous forme de corpus des produits directs et “palpables” de l'enquête empirique de terrain » (De Sardan, 2008, p. 118). L'observation du

terrain « en ligne » permet pour sa part de contextualiser les données produites sur les dispositifs qui sont issues de l'extraction systématique. Concrètement, les données ici prennent deux formes : (1) des copies d'écran dans le cadre du travail d'observation ; (2) des tableaux Excel comportant notamment l'ensemble des messages publiés sur le dispositif Twitter par les acteurs au cœur de l'enquête – entre autres, les parlementaires européens, porte-paroles, journalistes et responsables de la communication.

Si nous proposons ci-dessous d'indiquer les étapes significatives de l'analyse, il ne faudrait pas pour autant croire que celles-ci se déclinent de manière linéaire. La démarche est itérative, certaines activités se déroulent en parallèle, d'autres successivement, avec des allers-retours entre les différents types de données.

La mise en ordre

La mise en ordre, autrement dit le classement des données, est une étape essentielle dans l'entreprise de « mise en sens » des données. Si celles-ci n'entrent pas dans la catégorie des données massives, leur addition produit néanmoins des corpus de grande quantité réunissant des matériaux de format très hétérogène. Cela n'est pas sans poser un certain nombre de défis. Plusieurs stratégies ont été choisies pour inventorier et classer les données. Dans un dossier nommé « corpus », des sous-dossiers ont été créés : (1) par thématique ; (2) par acteur ; (3) par séquence temporelle ; (4) par type de données ; et le cas échéant ; (5) par événement particulier. Un sous-dossier constitué des (6) sources documentaires (documents institutionnels, etc.) a également été créé. Il est par conséquent possible que des doublons de données (comme des captures d'écran, des fichiers Excel ou des transcriptions d'entretiens, etc.) se retrouvent dans différents dossiers. Réaliser le travail de classement a permis de gagner en connaissance et en familiarité avec le corpus de données.

La mise à plat

Si les captures d'écran sont stockées dans leur format initial, les autres types de données ont généralement été imprimés pour pouvoir les analyser : les notes d'observation, les transcriptions d'entretiens, mais aussi les messages

publiés (*tweets*). Ces derniers étant difficiles à analyser finement dans le format Excel, nous avons préféré réduire le corpus qu'ils constituent. Par exemple, pour l'analyse des contenus produits sur les dispositifs par les candidates et candidats aux élections européennes de mai 2009, nous avons établi une liste de critères permettant la sélection d'un nombre limité de comptes : des critères sociotechniques (comme le nombre de messages, nombre d'abonnés, nombre d'abonnements, etc.) et des critères sociopolitiques, comme la place sur la liste en période d'élection, les rôles institutionnels tenus pour les députés sortants, etc. (Roginsky et De Cock, 2015). Il est ici difficile de quantifier le nombre de pages que le corpus regroupant ces différents matériaux génère. Celles-ci ont d'abord fait l'objet de plusieurs lectures flottantes, pour s'imprégner du matériau, puis davantage focalisées – en fonction des objectifs poursuivis (autrement dit des questions de recherche).

La mise en lien

La démarche étant itérative, la mise en lien entre les données se réalise au fur et à mesure du travail d'analyse, dès la phase de classement des données. Le travail de mise en lien permet de faire dialoguer l'ensemble des données. Il nécessite une démarche de va-et-vient permanente pour permettre de lire les données issues de la collecte « en ligne » à la lumière des données issues de la collecte « hors-ligne », et vice versa.

Le travail de mise en lien est d'autant plus important qu'il va de pair avec un travail de contextualisation, qui est le « principal “garde-fou” contre toutes les généralisations hâtives ou indues » (Beaud et Weber, 2010, p. 217). La contextualisation n'est pas seulement sociologique, mais également sociotechnique afin de prendre en compte la nature et le rôle du dispositif dans la production discursive et picturale. Le travail de mise en lien est également important, car, la recherche se déroulant sur un temps long, il est nécessaire de toujours tenir compte de la séquence temporelle (les propos tenus au sujet des dispositifs siconumériques en 2010 sont nécessairement différents de ceux tenus en 2016, par exemple).

La mise en écriture

« C'est au moment de la rédaction que [l']analyse prend corps » (Beaud et Weber, 2010, p. 227) et permet d'assembler les différentes pièces du puzzle.

Les concepts viennent alors éclairer le terrain, et réciproquement. Ici, il est indispensable de revenir à la question de recherche: c'est elle qui va déterminer le type d'assemblage à réaliser. Quand nous nous sommes intéressées à la présentation de soi des candidats et candidates aux élections européennes sur Twitter, nous avons approché nos données en privilégiant la collection extensive de *tweets* pour avoir accès aux productions discursives publiées sur le dispositif, les captures d'écran récoltées lors de l'observation et les propos des personnes interrogées s'exprimant sur cet aspect particulier (Roginsky et De Cock, 2015). Quand nous nous sommes intéressées à la fabrique de la communication des membres du Parlement européen, nous sommes parties des données récoltées « hors-ligne » en intégrant la nature des dispositifs (Roginsky et Jeanne-Perrier, 2014; Roginsky, 2015). Quand nous nous sommes intéressées à leurs publics sur Facebook et Twitter, les données issues des entretiens ont été essentielles, mais elles ont été complétées et éclairées par des données issues de l'observation « en ligne » (Roginsky et Huys, 2015).

Le bricolage et la triangulation

L'enquête de terrain procède du bricolage. Le bricolage, pour Denzin et Lincoln (1998), est un modèle de recherche sociale caractérisé par une flexibilité et un pragmatisme réels, qui permet de s'adapter à la situation à laquelle on fait face. Il faut par conséquent accomplir un nombre important de tâches différentes, qui vont de la conduite d'entretiens à l'observation, à l'interprétation de documents personnels, en passant par l'introspection et la réflexivité. La triangulation est ainsi l'outil favori du bricolage, puisqu'elle permet de collecter différents types de matériaux. Des matériaux de nature disparate sont assemblés pour produire une construction émergente, une « création » complexe, dense et réflexive⁵. Une telle démarche induit un mouvement de va-et-vient pour assembler les différents matériaux. C'est aussi un mouvement entre le niveau micro (*zoom-in*) et un niveau un peu plus macro (*zoom-out*) pour cerner les « schèmes socialement pertinents » (Lahire, 2011, p. 142) dans la situation étudiée, ce qui pose la question du contexte et de son appréciation dans la recherche.

5. Voir Markham (2005) pour un exemple de présentation de travail de bricolage rassemblant dans un même article des matériaux hétérogènes.

Selon Marwick et boyd (2011), le contexte disparaît sur les dispositifs socionumériques (ces dernières parlent de *context collapse*), du moins aux yeux des publics. Si, effectivement, les traces numériques « ont tendance à effacer le contexte social et biographique dans lequel s'inscrivent les individus pour se concentrer sur une forme particulière de l'activité » (Ouakrat et Mésengeau, 2016), ce contexte n'en demeure pas moins un élément compréhensif important. La triangulation des méthodes et des données permet à ce titre de ne pas chercher à aplatir, voire à gommer une complexité inhérente au phénomène étudié, mais de tenter au contraire de l'appréhender au plus près.

L'ethnographie joue dans ce cadre un rôle particulier, car elle permet de produire des données épaisses, détaillées et denses (Geertz, 1973), interconnectées (Stake, 1985), riches (Becker, 1970) et contextualisées. Le bricolage est un élément essentiel de la démarche ethnographique puisque celle-ci nécessite de s'adapter aux terrains étudiés, aux multiples points de vue pour les approcher, en ajustant les méthodes de collecte et d'analyse de données en conséquence. L'ethnographie donne une place importante aux représentations des acteurs étudiés, élément indispensable à la compréhension du contexte social. Celles-ci ne sont pas extrapolables des contenus produits en ligne. Elles sont généralement liées à la compréhension que les individus ont de leur contexte d'action. D'où la nécessité d'appréhender les contextes au plus près des expériences des individus.

La démarche ethnographique incite également à réfléchir à son propre rôle et à son positionnement sur le terrain étudié, en tant que chercheuse et chercheur. Ici encore le bricolage est utile, puisqu'il permet, selon Wibberley (traduction libre de 2012, p. 5), de réaliser un « produit construit consciemment ». Les commentaires réflexifs deviennent le ciment qui permet d'assembler les différents matériaux collectés. Le bricolage a ainsi le mérite de mettre en lumière la dimension interprétative de la recherche, et par voie de conséquence, la manière dont on participe à la construction de la recherche.

La triangulation permet également d'éviter le traitement isolé des données de différentes natures et, au contraire, de les enrichir en les faisant dialoguer. La question du statut des données n'est pas pour autant réglée : comment savoir si les données recueillies sont publiques ou privées ? Si le statut des données recueillies directement auprès des acteurs peut être qualifié avec eux (en assurant par exemple l'anonymat des propos, comme

c'est généralement le cas dans notre recherche), qu'en est-il des données recueillies à leur insu ? Si la question se pose pour les données récoltées dans le cadre d'observations « hors-ligne », elle prend davantage d'acuité encore avec les données récoltées « en ligne ». Par exemple, peut-on considérer que les données publiées « en ligne » sont des indicateurs de la nature publique des activités « hors-ligne » dont elles font le récit ? Le fait d'être visible « en ligne » induit-il automatiquement la visibilité de l'activité « hors-ligne » ? L'Association of Internet Researchers a produit à cet égard des recommandations éthiques⁶ et elle renvoie à la notion d'« intégrité contextuelle » (*contextual integrity*) de Nissebaum (2004) pour qualifier la nature des données. Le caractère public ou privé des données dépendrait ainsi du contexte spécifique, où l'on retrouve une fois encore la notion de contexte.

* * *

La triangulation proposée dans ce travail de recherche sur les usages des dispositifs socionumériques par les personnels politiques s'ancre dans un paradigme qualitatif. Il nous semble en effet que seules les approches qualitatives peuvent aller au-delà des traces d'usage pour explorer véritablement les usages comme des pratiques sociales. Granjon (2014) nous engage ainsi à ne pas prendre les premières pour les secondes, autrement dit à ne pas prendre pour explication ce qui doit être expliqué. En effet, les traces numériques ne peuvent pas nous permettre, à elles seules, de saisir la complexité du contexte social et de comprendre « l'ensemble de l'activité d'un individu » (Ouakrat et Mésengeau, 2016).

Pour analyser ces usages professionnels, une profonde connaissance du contexte des individus étudiés est indispensable, en partant de la compréhension que ceux-ci en ont. Une telle démarche permet d'éviter les risques de surinterprétation, que De Sardan (1996, p. 31) définit comme « un excès de sens imposé au "réel" dont on prétend rendre compte », qui se présentent particulièrement quand on se limite à l'analyse des données collectées sur le terrain en ligne. Si les approches ethnographiques sont à cet égard particulièrement utiles, car elles permettent de resituer l'ancrage social des pratiques, elles ne constituent pas pour autant la panacée :

6. Voir aoir.org/reports/ethics2.pdf

au-delà des difficultés pratiques qu'elles peuvent générer (accessibilité aux terrains de recherche et aux individus, statut des propos d'entretien, etc.), elles peuvent amener à surévaluer le poids des contextes sociaux ou à adhérer, avec une trop grande proximité, aux propos recueillis auprès des acteurs étudiés. D'autres méthodes de collecte et d'analyse des données peuvent alors être employées, comme l'observation des activités qui se déclinent sur le terrain « hors-ligne », l'approche sémiologique, qui permet d'analyser finement les dispositifs sociotechniques, mais aussi l'extraction automatique de données produites sur les dispositifs. Il est ainsi possible de mettre en lien les traces d'activités numériques avec le monde social.

Références

- BEAUD, Stéphane et Florence WEBER, *Guide de l'enquête de terrain*, La Découverte, 2010.
- BECKER, Howard S., *Sociological Work: Method and Substance*, Transaction Publishers, 1970.
- BONACCORSI, Julia, « Approches sémiologiques du Web », dans Christine BARATS (dir.), *Manuel d'analyse du Web en sciences humaines et sociales*, Armand Colin, 2013, p. 125-146.
- CHAMBOREDON, Hélène *et al.*, « S'imposer aux imposants. À propos de quelques obstacles rencontrés par des sociologues débutants dans la pratique et l'usage de l'entretien », *Genèses*, vol. 16, 1994, p. 114-132.
- COHEN, Louis et Lawrence MANION, *Research Methods in Education*, Croom Helm, 1986.
- COUTANT, Alexandre et Thomas STENGER, « L'avènement des médias sociaux ? Une histoire de participation », *Le Temps des médias*, vol. 1, n° 18, 2012, p. 76-86.
- DE SARDAN, Jean-Pierre, *La rigueur du qualitatif. Les contraintes empiriques de l'interprétation socio-anthropologie*, L'Harmattan, 2008.
- DE SARDAN, Jean-Pierre, « L'enquête de terrain socio-anthropologique », dans Jean BOUTIER, Jean-Louis FABIANI et Olivier DE SARDAN (dir.), *Corpus, sources et archives*, Institut de recherche sur le Maghreb contemporain, 2001, p. 63-81.
- DE SARDAN, Jean-Pierre, « La violence faite aux données. De quelques figures de la surinterprétation en anthropologie », *Enquête*, n° 3, 1996, p. 31-59.
- DENZIN, Norman, K. et Yvonna S. LINCOLN, *Strategies of Qualitative Inquiry*, SAGE Publications, 1998.
- DUMONT, Guillaume, « Entre mobilité, virtualité et professionnalisation : éléments méthodologiques et conditions de réalisation d'une ethnographie des grimpeurs professionnels », *Recherches qualitatives*, vol. 33, n° 1, 2014, p. 188-210.
- FOURNIER, Pierre, « Observation », dans Serge PAUGAM (dir.), *Les 100 mots de la sociologie*, Presses universitaires de France, 2010, p. 36-37.

- GEERTZ, Clifford, *The Interpretation of Cultures: Selected Essays*, Basic books, 1973.
- GLASER, Barney G. et Anselm L. STRAUSS, *The Discovery of Grounded Theory: Strategies for Qualitative Research*, Aldine, 1967.
- GRANJON, Fabien, « Problematizing social uses of information and communication technology: A critical French perspective », *Canadian Journal of Communication*, vol. 39, n° 1, 2014, p. 111-125.
- JOUËT, Josiane, « Usages et pratiques des nouveaux outils de communication », dans Lucien SFEZ (dir.), *Dictionnaire critique de la communication*, Presses universitaires de France, 1993.
- KAVANAGH, Éric, Jacynthe ROBERGE et Isabelle SPERANO, « Typologie exploratoire des affordances textuelles », *Pratiques. Linguistique, littérature, didactique*, p. 171-172, 2016.
- LAHIRE, Bernard, *L'homme pluriel. Les ressorts de l'action*, Pluriel, 2011.
- LAURENS, Sylvain, « “Pourquoi” et “comment” poser les questions qui fâchent ? Réflexions sur les dilemmes récurrents que posent les entretiens avec des “imposants” », *Genèses*, vol. 4, n° 69, 2007, p. 112-127.
- LEANDER, Kevin M. et Kelly MCKIM, « Tracing the everyday “sittings” of adolescents on the Internet: A strategic adaptation of ethnography across online and offline spaces », *Education, Communication & Information*, vol. 3, n° 2, 2003, p. 211-240.
- MABI, Clément, « Décrypter le *design* des technologies numériques. Un enjeu pour une culture numérique plus réflexive », *Hybrid: Revue des arts et médiations humaines*, n° 3, 2016.
- MARKHAM, Annette, « “Go ugly early” : Fragmented narrative and bricolage as interpretive method », *Qualitative Inquiry*, vol. 11, n° 6, 2005, p. 813-839.
- MARWICK, Alice E. et danah BOYD, « I tweet honestly, I tweet passionately: Twitter users, context collapse and the imagined audience », *New Media & Society*, vol. 13, n° 1, 2011, p. 114-133.
- MERCKLÉ, Pierre, « Corpus », dans Serge PAUGAM (dir.), *Les 100 mots de la sociologie*, Presses universitaires de France, 2010, p. 28-29.
- MILLERAND, Florence, « Usages des NTIC: Les approches de la diffusion, de l'innovation et de l'appropriation », *COMMposite*, vol. 98.1., 1998.
- NISSEBAUM, Helen, « Privacy as contextual integrity », *Washington Law Review*, vol. 79, n° 119, 2004, p. 101-139.
- OUKRAT, Alain et Julien MÉSENGEAU, « Resocialiser les traces d'activités numériques: une proposition qualitative pour les SIC », *Revue française des sciences de l'information et de la communication*, n° 8, 2016.
- ROGINSKY, Sandrine, « Les députés européens sur Facebook et Twitter: une ethnographie des usages », *Communication & Langages*, n° 183, 2015, p. 83-109.
- ROGINSKY, Sandrine et Barbara DE COCK, « Faire campagne sur Twitter. Modalités d'énonciation et mises en récit des candidats à l'élection européen », *Les Cahiers du Numérique*, vol. 11, n° 4, 2015, p. 119-144.
- ROGINSKY, Sandrine et Sophie HUYS, « À qui parlent les professionnels politiques ? La fabrique des publics des députés européens sur les dispositifs Facebook et Twitter », *Communication*, vol. 33, n° 2, 2015.

- ROGINSKY, Sandrine et Valérie JEANNE-PERRIER, « La fabrique de la communication des parlementaires européens. « Tweet ton député » et les « Ateliers du député 2.0 », *Politiques de communication*, vol. 2, n° 3, 2014, p. 85-124.
- SOULÉ, Bastien, « Observation participante ou participation observante? Usages et justifications de la notion de participation observante en sciences sociales », *Recherches qualitatives*, vol. 27, n° 1, 2007, p. 127-140.
- STAKE, Robert, E., « Case study », dans John NISBET, Jacquetta MEGARRY et Stanley NISBET (dir.), *World Yearbook of Education 1985: Research, Policy and Practice*, Kogan Page, 1985, p. 277-285.
- STECK, Jean-Fabien, « Être sur le terrain, faire terrain », *Hypothèses*, n° 15, 2012, p. 75-84.
- VAN DIJCK, José, *The Culture of Connectivity: A Critical History of Social Media*, Oxford University Press, 2013.
- WIBBERLEY, Christopher, « Getting to grips with bricolage: A personal account », *The Qualitative Report*, vol. 17, n° 50, 2012, p. 1-8.

CHAPITRE 7

Les pratiques de visionnement connecté des jeunes : méthodes traditionnelles, virtuelles et numériques

Christine Thoër, Florence Millerand, Nina Duque

Le visionnement de contenus vidéo (émissions de télé-réalité, talk-shows, séries télévisuelles, webséries, capsules vidéo) se fait de plus en plus de manière connectée et cette tendance touche particulièrement les jeunes (CEFRIQ, 2017a ; Kervella et Loicq, 2015). Plusieurs facteurs contribuent à cette tendance, entre autres, l'amélioration des connexions réseau, la progression en équipements mobiles (téléphone intelligent) et le développement de l'offre (pensons à Netflix). Nous avons mené une recherche auprès de jeunes âgés de 12 à 25 ans résidant au Québec, dans le but de cerner en quoi ce nouvel univers de consommation, marqué à la fois par l'usage de nouveaux dispositifs techniques et par une augmentation et une diversification de l'offre de contenus, transforme l'expérience de visionnement et l'attachement aux contenus¹. Dans ce chapitre, nous abordons les aspects méthodologiques liés à l'étude de ces pratiques et nous discutons l'apport combiné de différents types de méthodes en soulignant les principaux défis et enjeux associés.

Cerner les pratiques des jeunes en matière de visionnement connecté implique de déployer des stratégies méthodologiques aptes à les saisir

1. La présentation détaillée du projet « Jeunes et visionnement connecté, quelles mutations des cultures médiatiques » est disponible en ligne : blogsgrms.com/jeunes-visionnement-connecte/

dans toute leur diversité et complexité. D'une part, ces pratiques apparaissent de plus en plus morcelées et révèlent une tendance à l'hyper-personnalisation (Kervella et Loicq, 2015; Thoër *et al.*, 2015). D'autre part, elles s'entremêlent à un ensemble d'autres activités, tant en ligne que hors-ligne, tout en prenant place dans une diversité d'espaces et de temporalités. En outre, elles évoluent rapidement, au fur et à mesure de l'autonomisation des jeunes et en parallèle de changements continus dans l'écosystème médiatique.

Un portrait du paysage méthodologique actuel comprend différentes orientations aptes à appréhender des phénomènes ou des pratiques sociales intervenant en ligne (voir Millerand, Myles et Proulx, le présent ouvrage, pour une critique du portrait brossé par Marres, 2012). Parmi ces orientations méthodologiques, les méthodes dites traditionnelles (ou *methods-as-usual*, selon l'expression de Marres (2012), regroupent les méthodes de recherche classiques des sciences sociales (habituellement distinguées en fonction de leur orientation qualitative, quantitative ou mixte), qu'il s'agit de transposer dans le contexte numérique. La recherche dont nous rendons compte dans ce chapitre relève de cette catégorie. Les méthodes dites virtuelles, popularisées par les travaux de Hine (2017) sur l'ethnographie virtuelle, mobilisent quant à elles une perspective ou des outils ethnographiques pour étudier des phénomènes socioculturels prenant place en ligne dans une tradition socio-anthropologique (voir Hine, le présent ouvrage). Enfin, les méthodes dites numériques (ou *digital methods*, selon l'expression de Rogers, 2013) sont utilisées elles aussi pour appréhender des phénomènes sociaux tels qu'ils se manifestent en ligne, mais en ayant recours à des méthodes et des outils natifs des plateformes du numérique (comme les données fournies par Google Analytics, par exemple)².

Dans le cadre de ce chapitre, nous reprenons ces grandes catégories de méthodes pour apporter des éléments de réponse à la question suivante : jusqu'où peut-on aller dans la compréhension des pratiques de visionnement connecté chez les jeunes en employant des méthodes traditionnelles, et quelles autres méthodes peuvent être mobilisées pour pallier leurs limites éventuelles ? Nous commençons en présentant rapi-

2. Précisons que Marres (2012) ajoute à ces trois grandes orientations méthodologiques les méthodes dites computationnelles (ou *big methods*) associées à l'utilisation de techniques informatiques d'extraction, de gestion, d'analyse et de visualisation de vastes corpus de données (*big data*) issues notamment du Web et des médias sociaux.

dement les méthodologies employées dans les études de réception. Nous présentons ensuite la stratégie méthodologique de la recherche que nous avons menée sur les pratiques de visionnement connecté de jeunes adultes. Puis, nous discutons des forces et des limites des méthodes dites traditionnelles pour l'étude des pratiques de visionnement connecté chez les jeunes, pour aborder finalement l'apport combiné de méthodes virtuelles et numériques à l'analyse des pratiques de visionnement connecté.

Les méthodes utilisées dans les études empiriques de réception

En s'attachant à ne pas documenter uniquement le travail d'interprétation des contenus (ou des « textes » dans le modèle récepteur/lecteur), les travaux sur la réception visent la compréhension de ce que « font » les individus lorsqu'ils regardent la télévision. Les travaux ont porté sur les contextes de réception, les cadres spatio-temporels de l'écoute et les rapports sociaux qui les traversent pour montrer leur rôle dans la construction de significations autour de la pratique télévisuelle (Corner, 1996; Silverstone *et al.*, 1991). Ils ont permis de documenter comment l'activité télévisuelle s'inscrit à la fois dans l'espace domestique, en structurant les rituels quotidiens et les identités familiales et de genre (voir Morley, 1986; Proulx et Laberge, 1995) et dans l'espace social, en permettant de se lier grâce au partage de répertoires d'actions et de références communes (voir Boullier, 2004). Ces travaux ont marqué le « tournant ethnographique » de la réception en s'appuyant sur une multitude de méthodes : immersion dans les foyers, observation en contexte, entrevues en profondeur, récits de pratiques, recours à des carnets de temps (ou journaux d'écoute), journaux personnels, listes d'équipements, dessins et photographies des installations.

Plus récemment, les recherches ont investi les espaces en ligne pour comprendre leur articulation avec les pratiques de réception. Certains travaux se sont intéressés à documenter le déploiement des discours et des conversations sur les réseaux sociaux numériques autour de la télévision (Spies, 2016) ou des films (Allard, 2000). Dans la lignée des études sur le courrier des lecteurs (Pasquier, 1999), des travaux ont été menés sur les communautés constituées autour de séries télévisées ou de films, mettant en évidence les activités créatrices, collaboratives et de sociabilité des « fans » (Bourdaa, 2015), pour certains dans une perspective autoethnographique (Jenkins,

1992). Plus récemment les formes de la critique amateur sur les plateformes en ligne ont fait l'objet d'études mobilisant des méthodologies diversifiées qui montrent comment ces contributions sont utilisées pour produire des classifications, des évaluations et des recommandations permettant de valoriser les contenus culturels (Béliard, 2016 ; Pasquier *et al.*, 2014). Parallèlement à ces études principalement qualitatives se sont développées des études par questionnaire, qui, outre la mesure des audiences télévisuelles par des panels d'audimètre, visent à repérer les déterminants des pratiques télévisuelles et plus généralement des pratiques culturelles (Gire et Granjon, 2012).

Aujourd'hui, le développement et la convergence des technologies incitent à réinterroger les pratiques de consommation de la télévision et plus largement des contenus culturels visionnés en ligne, ainsi que les rapports sociaux qui les entourent. Il semble en effet qu'un contenu n'est pas regardé tout à fait de la même façon et qu'il ne s'apprécie pas selon les mêmes termes lorsqu'il se donne à voir sur un écran d'ordinateur portable (ou de téléphone mobile), entre autres parce que les dispositifs connectés ouvrent à d'autres contextes de visionnement et affectent le processus d'interprétation et d'attachement aux contenus. Sur ce plan, la sociologie des usages a permis d'enrichir les cadres d'analyse en ajoutant la dimension du rapport tangible aux dispositifs techniques, afin de cerner en quoi ceux-ci sont porteurs « d'effets de sens » (Combes, 2013 ; Millerand *et al.*, 2016 ; Thoër *et al.*, 2015).

À ce jour, les études sur les pratiques télévisuelles en contexte connecté s'appuient sur les méthodes privilégiées par les études de réception de la télévision mobilisant des démarches qualitatives ou mixtes (entrevue, observation participante, enquête, carnet de temps), qui font généralement l'objet d'une triangulation, pour obtenir une meilleure compréhension des pratiques des téléspectateurs (voir par exemple Kervalla et Loicq, 2015). Ces méthodologies, qui consistent à transposer les méthodes classiques correspondant aux méthodes dites traditionnelles, présentent des atouts, mais aussi certaines limites, entre autres parce que la multiplication des dispositifs entraîne un morcellement des pratiques qui s'avèrent dès lors plus difficiles à circonscrire (Livingstone et Local, 2017). L'objectif de ce chapitre est précisément de reconnaître les enjeux méthodologiques que posent ces nouveaux contextes de visionnement, qui restent encore peu examinés et discutés.

Une étude des pratiques de visionnement connecté chez les jeunes au Québec

Si les travaux analysant l'évolution des pratiques télévisuelles dans l'écosystème médiatique actuel ont sans surprise porté surtout sur les contenus issus de la télévision, les pratiques des jeunes en ligne impliquent, outre les contenus télé, d'autres contenus produits par une diversité d'acteurs et circulant notamment sur YouTube et sur les réseaux sociaux numériques (Thoër *et al.*, 2015). Pour comprendre comment la coexistence de ces contenus transforme la façon dont les jeunes les découvrent, les regardent, s'y attachent et les partagent, nous avons cherché à documenter l'ensemble des pratiques de visionnement connecté des jeunes, dès lors qu'elles s'inscrivaient dans une perspective de divertissement. La recherche visait ainsi à repérer les contenus regardés, les catégorisations utilisées, les contextes de visionnement (dispositifs, lieux et installations, temps de visionnement, activités connexes), ainsi que les significations associées à ces pratiques et leur inscription dans les sociabilités juvéniles.

Pour ce faire, nous avons déployé une stratégie méthodologique s'appuyant sur une diversité de techniques de collecte qualitatives et quantitatives (groupes de discussion, entrevues, carnets de temps remplis en ligne, enquêtes par questionnaire), permettant une triangulation des résultats. La collecte s'est déroulée sur trois ans, de 2014 à 2017. Privilégiant une approche partenariale, nous avons tout d'abord conduit deux groupes de discussion rassemblant des représentants d'entreprises de production et d'associations médiatiques ainsi que des diffuseurs, experts et régulateurs, afin de cerner leurs préoccupations et de mieux anticiper les retombées du projet.

Par la suite, dix groupes de discussion d'une durée d'une heure trente à deux heures ont été réalisés en 2015 auprès de jeunes âgés de 12 à 25 ans résidant au Québec (répartis selon le genre et en quatre tranches d'âge: 12-13 ans, 14-16 ans, 17-19 et 20-25 ans). Nous avons rencontré un total de 61 jeunes (28 filles et 33 garçons) vivant à Montréal, d'origine culturelle variée, qui fréquentaient des établissements d'enseignement en français, certains et certaines, parmi les plus âgés, étant exclusivement sur le marché du travail. Le recrutement s'est fait de plusieurs façons: par réseaux de connaissances, dans une école secondaire publique et une maison de jeunes, par des annonces en ligne et par la méthode dite de la

« boule de neige »³. Le critère de recrutement que dépassaient très largement tous les jeunes était de regarder au moins deux contenus de divertissement en ligne par semaine.

Fin 2015 et début 2016, nous avons réalisé des entretiens avec 29 jeunes (17 filles et 12 garçons), recrutés pour la plupart parmi ceux et celles ayant participé aux groupes de discussion. Ces entrevues visaient à documenter les pratiques personnelles des jeunes en matière de visionnement connecté et à cerner leur évolution au cours de la dernière année. L'objectif était de saisir l'activité réflexive qu'exercent les jeunes sur leurs pratiques.

Dans le cadre des entrevues, nous avons envoyé aux volontaires un carnet de temps à remplir en ligne afin de prendre connaissance dans le détail de tous les contenus audiovisuels visionnés au quotidien. Ce carnet a été rempli par 12 des 29 jeunes (dont 10 filles). Tous ceux ayant participé aux groupes de discussion et aux entrevues ont rempli un questionnaire visant à recueillir des données de type sociodémographique ainsi que des informations sur les équipements utilisés et le temps consacré au visionnement connecté.

Enfin, en 2017, nous avons réalisé en partenariat avec trois diffuseurs télévisuels et le CEFRIO (2017b) une enquête en ligne auprès d'un échantillon (n=1504) tiré d'un panel d'internautes. Cette dernière phase de collecte visait à quantifier les tendances observées et à cerner les variations des pratiques de visionnement selon l'âge, le revenu, le contexte familial et l'équipement.

Les apports et les limites des méthodes traditionnelles

Les méthodes traditionnelles se sont révélées particulièrement utiles pour dresser le portrait des contenus visionnés par les jeunes. En revanche, elles ont montré certaines limites dans leur capacité à saisir la pluralité des contextes de visionnement.

3. Dans le cadre de cette méthode, on demande aux participants de désigner des personnes de leur réseau qui correspondent aux critères de recrutement et qui sont invitées à participer à la recherche.

Cerner la diversité des contenus et les catégories les désignant

Les groupes de discussion ont permis de désigner les types de contenus visionnés par les jeunes et de cerner leur manière d'en parler. Ces jeunes rapportaient spontanément les titres et les noms des acteurs et actrices, animateurs et animatrices d'émissions ou de YouTubeurs qu'ils suivaient. En stimulant par entraînement la mémorisation et l'évocation des contenus, la technique du groupe de discussion s'est révélée être un mode de collecte particulièrement approprié. La dynamique de groupe permettait aussi d'accéder aux contenus qui semblaient faire consensus (par exemple, la série *How I Met Your Mother*, citée dans tous les groupes). De plus, le moindre contrôle exercé par l'animatrice et l'effet d'entraînement sur les jeunes ont permis, par exemple aux garçons âgés de 17 à 25 ans, de dépasser les tabous concernant les contenus pornographiques (très peu mentionnés en entrevue individuelle). À l'inverse, certains contenus considérés comme moins légitimes peuvent ne pas avoir été mentionnés, les entrevues restant le moyen privilégié pour cerner les pratiques individuelles.

Si les groupes de discussion ont permis de discerner les types de contenus visionnés, les jeunes restaient perplexes lorsque nous leur demandions de les catégoriser. Ils ne jugeaient pas nécessaire de nommer ces contenus pour les trouver et les regarder, distinguant surtout entre formats longs (séries et films) et courts (vidéos) (Thoër *et al.*, 2015). Aussi, l'exercice était-il un peu forcé même si certaines catégories ont émergé et semblaient s'imposer, du moins au sein de certains groupes, pour les vidéos courtes (*tuto*, *vlog*, *fail*, par exemple) et les centres d'intérêt auxquels elles sont associées (par exemple, *make-up*, *gaming*).

Rendre compte de la pluralité des contextes du visionnement connecté

La rapidité avec laquelle l'écosystème médiatique contemporain évolue complique l'étude des pratiques de visionnement connecté, ces mutations étant elles-mêmes difficiles à saisir au fil du temps. Parallèlement à une collecte de données qui s'est déroulée sur trois ans, nous avons documenté l'évolution de l'écosystème médiatique grâce à une veille de l'actualité technique, économique, juridique et politique en lien avec le secteur du divertissement connecté. Celle-ci nous a permis de prendre connaissance, entre autres, des formules de vidéo à la demande disponibles au Québec.

En un an, nous avons pu observer la baisse de l'utilisation des plateformes de *streaming* (en continu) illégal au profit du service Netflix, apprécié par les jeunes pour sa convivialité, comme cela s'est produit avec les plateformes d'écoute de musique en continu comme Deezer ou Spotify. Toutefois, il n'est pas toujours facile de départager ce qui relève de l'évolution de l'offre, de l'expérience ou de l'âge des jeunes dans la transformation des pratiques.

En contexte connecté, les contenus de divertissement sont accessibles grâce à une multitude d'écrans, et leur visionnement s'insère dans différentes activités et pratiques de communication. Les jeunes peinaient à décrire les plateformes utilisées pour accéder aux contenus, de même que les stratégies et dispositifs sur lesquels ils s'appuyaient pour en découvrir de nouveaux. Ils développent en effet des routines de visionnement qui ne sont pas toujours réfléchies et sont souvent fondées sur la sérendipité et les propositions algorithmiques. De plus, les jeunes ne cherchent pas à mémoriser les noms ou les adresses URL des plateformes, à l'exception de certaines d'entre elles comme Netflix, parce qu'ils procèdent par la mise en favori et surtout parce qu'ils accèdent aux contenus en effectuant une recherche dans les moteurs de recherche. Les entrevues individuelles ou de groupe s'avèrent ainsi peu adaptées pour appréhender la construction et la stabilisation de ces routines de visionnement. Dans le cadre de cette recherche, nous avons contourné cet enjeu en procédant à une reconstitution des routines de visionnement des jeunes en entrant à notre tour les contenus nommés dans les moteurs de recherche. Cette technique nous permettait de mieux comprendre les processus de recherche de contenus des jeunes et surtout de savoir à quoi ces contenus se rapportaient. Dans le cadre des entrevues individuelles, nous invitons aussi les jeunes à nous montrer certains des contenus qu'ils visionnaient ainsi que les plateformes utilisées afin d'ancrer leurs récits dans des exemples concrets.

Par ailleurs, il s'est avéré difficile d'amener les jeunes à décrire leurs installations de visionnement. Nous leur avons demandé d'apporter des photos de ces installations pour enrichir notre compréhension, mais rares sont ceux qui l'ont fait. Cette pratique de photographie semble plus évidente si c'est la chercheuse qui l'assume, par exemple, dans le cadre d'entrevues au domicile (Blanc, 2015). En outre, il n'était pas évident de cerner le temps dédié au visionnement connecté. Seuls 12 jeunes ayant

participé aux entrevues avaient complété le carnet de temps, malgré nos efforts pour le rendre convivial et l'octroi d'un dédommagement sous forme de bon d'achat.

Les récits des jeunes montrent une consommation fragmentée où les séries jouxtent les émissions de cuisine, les vidéos de YouTubeurs, les émissions de télé-réalité musicale et les clips de musique (pour ne nommer que ces contenus), sans compter les détours par les réseaux socionumériques. Le format court des vidéos, ainsi que leur visionnement à l'aide des périphériques mobiles, font éclater l'association classique entre pratique et support en permettant à la fois (et souvent en même temps) d'écouter de la musique, de regarder un film, d'écrire, de faire ses devoirs, de manger, de faire le ménage, etc. Au final, le visionnement se superposant à un ensemble d'autres activités et les pratiques prenant place dans une diversité d'espaces (chez soi, à l'école ou à l'université, au travail, dans les transports) et de temporalités, déterminer ce qui constitue l'activité principale ou du moins la plus significative reste un défi. Ces enjeux appellent à un réaménagement méthodologique, notamment grâce à l'inclusion des méthodes virtuelles et numériques, dont nous présentons les potentialités dans la section qui suit.

L'apport des méthodes virtuelles et numériques

Les méthodes virtuelles et numériques peuvent compléter les méthodes traditionnelles en offrant à la fois des descriptions plus fines des routines de visionnement et la possibilité de les observer dans une diversité d'espaces et de temporalités.

Les méthodes virtuelles

Les méthodes dites virtuelles consistent à adapter les méthodes de recherche qualitatives classiques, notamment ethnographiques, à l'environnement numérique qui est envisagé dans son articulation avec les environnements hors-ligne. À ce titre, elles peuvent venir pallier certaines limites des méthodes traditionnelles. S'intéresser aux pratiques de visionnement connecté implique de prendre en compte la multiplicité des contextes dans lesquels ces pratiques prennent place et donc de diversifier les contextes d'observation. Si les observations *in situ* ou les entrevues

réalisées à domicile peuvent permettre d'accéder aux installations et de favoriser la mémorisation des pratiques et ainsi produire des interprétations plus fines des pratiques et des contextes d'usage, ces méthodes très intrusives soulèvent des enjeux éthiques importants en plus d'exiger de longs séjours qui peuvent être difficiles à organiser. En outre, ces méthodes ne permettent pas de saisir les pratiques de visionnement se déroulant en contexte de mobilité. Sur ce plan, de nouveaux outils ont été proposés dans les dernières années pour produire des données d'observation en temps réel, comme les lunettes caméra (Licoppe et Figeac, 2014). Ce type d'outil permet de produire des données du point de vue de la participante ou du participant, et ainsi de resituer les usages dans leur contexte de production et dimension subjective. Il reste toutefois à mieux documenter comment ce type de dispositif méthodologique affecte l'expérience de visionnement, sans compter les nouveaux problèmes éthiques éventuels que peut soulever son utilisation.

Pour mieux comprendre comment se construisent les routines et cerner la signification des pratiques de visionnement connecté pour les individus, il pourrait être intéressant de procéder à des récits de pratique qui peuvent être réalisés en ligne. D'une part, cette méthode faciliterait le suivi d'une cohorte sur une plus longue période (Orgad, 2005) ; d'autre part, elle pourrait être couplée à la tenue d'un journal en ligne (blogue) qui permettrait aux jeunes de documenter leurs expériences de visionnement. Ceux-ci pourraient y produire des textes et des enregistrements audio ou vidéo (par exemple, avec leur téléphone), sur la façon dont ils découvrent et visionnent des contenus, tout en partageant des liens et des captures d'écran des sites et plateformes utilisés. Cette méthode dite du blogue privé, qui a déjà été utilisée en jumelage avec des entrevues semi-dirigées pour cerner la construction de l'agentivité sexuelle (Lang, 2016), présente plusieurs avantages permettant notamment de cerner des processus qui s'étalent dans la durée, de mieux comprendre le rôle du dispositif et de réduire l'asymétrie de la relation intervieweur-interviewé. Un tel dispositif pourrait ainsi favoriser l'expression sur les contenus visionnés moins légitimes, le processus narratif pouvant aussi activer les réflexions intimes que suscitent la pratique de visionnement ou les discussions qui l'entourent. À ce titre, il pourrait être judicieux d'utiliser des plateformes qui sont déjà familières aux jeunes publics, comme Instagram.

Par ailleurs, il pourrait être intéressant pour mieux cerner les ressources du Web qui sont utilisées dans le cadre des pratiques de visionnement connecté (recherche de contenus à visionner, recherche d'informations complémentaires sur les contenus visionnés, partage autour des contenus) de répertorier les ressources mentionnées par les jeunes en entrevue et de procéder à une ethnographie des espaces où les pratiques de visionnement connecté et les contenus qu'elles ciblent sont discutés. L'intérêt de ce type de méthodes est de donner accès à des données « passives », en ce sens qu'elles n'ont pas été construites par le dispositif de recherche (Pasquier *et al.*, 2014).

Les méthodes numériques

La collecte d'information sur l'ensemble des activités qui caractérisent le visionnement en ligne peut aussi se faire à partir des traces conservées dans l'historique de navigation, par exemple à partir d'outils issus directement de l'environnement numérique comme Google (pour cerner les modalités de découvrabilité des contenus) ou sur la plateforme Netflix (en utilisant les comptes personnels), ou encore par l'installation de logiciels d'enregistrement des traces d'activités. L'intérêt de ces méthodes dites numériques (Rogers, 2013; Snee *et al.*, 2016) est de procéder à des saisies plus systématiques et plus précises des parcours de divertissement en ligne que celles fondées sur les pratiques déclarées. Ce type de dispositif permettrait notamment de cerner avec plus de précision la diversité des ressources mobilisées dans le cadre du visionnement et le temps que les jeunes consacrent au visionnement de contenus en ligne, même si les modalités pour l'estimer restent attachées aux dispositifs techniques et non aux contenus visionnés eux-mêmes (Livingstone et Local, 2017).

Ces techniques prometteuses soulèvent toutefois de nouveaux défis, dont les difficultés de traitement et d'analyse de données parfois volumineuses ainsi que la nécessité d'installer ces outils sur une diversité d'appareils (téléphone, tablette, ordinateur), sachant que ces derniers sont souvent utilisés par plusieurs personnes au sein d'un même foyer et qu'il peut être difficile de détecter les traces d'un usager ou d'une usagère en particulier. Ces inconvénients peuvent être contournés partiellement en utilisant les traces (d'activités et d'usage de moteurs de recherche ou de plateformes de visionnement) comme matériau d'appui au processus

d'entrevue. La richesse de ces nouveaux types de données (captures vidéo des pratiques, traces d'activité) apparaît dès lors qu'elles sont soumises à la réflexivité des individus, appelés à les commenter, par exemple dans le cadre d'entrevues portant sur les traces (Dubois et Ford, 2015). Les productions réflexives des jeunes permettraient en effet de contextualiser leur parcours de visionnement, d'analyser les écarts possibles entre les traces captées et leurs perceptions et ainsi de saisir le sens que ces derniers attribuent à leurs pratiques et à ces écarts. Là encore, ces techniques posent des défis, à commencer par le nécessaire travail de sélection et de mise en scène des traces à commenter, impliquant la production de représentations et de visualisations aptes à susciter l'activité réflexive.

Notre recherche, qui mobilise une stratégie inspirée des méthodes utilisées dans les études de réception, confirme l'utilité d'une stratégie reposant sur la triangulation des méthodes éprouvées de la recherche qualitative (entre autres, groupes de discussion, entrevues, carnets de temps) et quantitative (enquêtes) et de leur transposition dans les environnements numériques. Elle montre l'importance d'envisager une approche à visée longitudinale pour saisir des pratiques en évolution, émergeant dans un contexte sociotechnique mouvant, auprès de populations jeunes dont les intérêts et les modes de vie se transforment rapidement. Les limites des méthodes employées, relevant des méthodes traditionnelles, témoignent notamment de difficultés à saisir l'évolution des parcours de visionnement connecté et la construction des routines, à cerner de manière fine et exhaustive la diversité des contenus visionnés et les usages des dispositifs connectés, à estimer les temps de visionnement et surtout à saisir comment ces pratiques s'inscrivent dans le quotidien et s'articulent à d'autres pratiques en ligne.

Les méthodes virtuelles qui tiennent compte des spécificités des terrains en ligne et de leur articulation avec les environnements hors-ligne, ainsi que les méthodes numériques qui s'appuient sur des outils issus de l'environnement numérique ou spécifiquement développés à des fins éducatives pour enregistrer les traces d'activité peuvent combler certaines limites des méthodes de recherche traditionnelles. Leur combinaison pour l'étude des pratiques de visionnement de contenus médiatiques en contexte connecté serait à explorer dans le cadre de travaux futurs. Pour cela, il serait important de cerner les modalités de mise en place de ces méthodes combinées et de recenser les ressources requises,

celles-ci semblant à première vue considérables (ne serait-ce que pour capturer et analyser adéquatement les traces laissées par les pratiques de visionnement connecté), et également de bien documenter les possibilités de triangulation des données.

Notre recherche témoigne enfin des difficultés à documenter les pratiques des jeunes issus de milieux défavorisés. En effet, si les résultats des groupes de discussion avaient révélé des usages légèrement différenciés chez les jeunes provenant de milieux plus populaires, ne serait-ce que sur le plan des équipements, l'enquête ne l'a pas complètement validé. Le fonctionnement des panels Web ne permet pas, malgré la pondération, d'assurer des échantillons représentatifs et tend à reproduire cette plus faible participation des milieux défavorisés (Vicente et Reis, 2016). L'accès aux usages des personnes issues de milieux défavorisés constitue un défi réel pour les études sur les usages connectés, car la fracture numérique qui semble persister, même dans les pays occidentaux, se joue aussi sur le plan de la participation aux recherches (Banaji, 2016 ; Pasquier, 2018).

Références

- ALLARD, Laurence, « Cinéphiles, à vos claviers ! Réception, public et cinéma », *Réseaux*, vol. 18, n° 99, 2000, p. 131-168.
- BANAJI, Shakuntala, *Global Research on Children's Online Experiences: Addressing Diversities and Inequalities*, Global Kids Online, Londres, 2016.
- BÉLIARD, Anne-Sophie, *S'il te plaît, dessine-moi une série. Contributions en ligne et séries télévisées*, Les Cahiers de l'EMNS, Presses des Mines, 2016.
- BLANC, Guillaume, « Les pratiques de réception télévisuelle dans les foyers à l'épreuve de l'audiovisuel numérique », *Études de communication*, n° 44, 2015, p. 63-78.
- BOULLIER, Dominique, *La télévision telle qu'on la parle: trois études ethnométhodologiques*, L'Harmattan, 2004.
- BOURDAA, Mélanie, « Les fans studies en question: perspectives et enjeux », *Revue française des sciences de l'information et de la communication*, vol. 7, 2015.
- CEFRIO, « Se divertir en ligne », *NETendances 2017*, vol. 8, n° 4, 2017a.
- CEFRIO, *Visionnement connecté par les jeunes de 12 à 25 ans au Québec*, 2017b.
- COMBES, Clément, *La pratique des séries télévisées: une sociologie de l'activité spectatorielle*, thèse de doctorat, ENMP, 2013.
- CORNER, John, « Sens, genre et contexte: la problématique du « savoir public » dans les nouvelles études de la réception », *Réseaux*, vol. 14, n° 79, 1996, p. 9-29.
- DUBOIS, Elizabeth et Heather FORD, « Trace interviews: An actor-centered approach », *International Journal of Communication*, vol. 9, 2015, p. 2067-2091.

- GIRE, Fabienne et Fabien GRANJON, « Les pratiques des écrans des jeunes français. Déterminants sociaux et pratiques culturelles associées », *RESET – Recherches en sciences sociales sur Internet*, vol. 1, 2012.
- HINE, Christine, « Ethnographies of online communities and social media: Modes, varieties, affordances », dans Nigel G. FIELDING, Raymond M. LEE et Garet BLANK (dir.), *The SAGE Handbook of Online Research Methods*, SAGE Publications, 2017, p. 401-413.
- JENKINS, Henry, *Textual Poachers: Television Fans and Participatory Culture*, Routledge, 1992.
- KERVILLA, Amandine et Marlène LOICQ, « Les pratiques télévisuelles des jeunes à l'ère du numérique: entre mutations et permanences », *Études de communication*, n° 44, 2015, p. 79-96.
- LANG, Marie-Ève, « Le blogue "privé" comme méthode de recherche en études féministes », *Recherches féministes*, vol. 29, n° 1, 2016, p. 71-90.
- LICOPPE, Christian et Julien FIGEAC, « L'organisation temporelle des engagements visuels dans des situations de multi-activité équipée en milieu urbain », *Activités*, vol. 11, n° 1, 2014, p. 65-85.
- LIVINGSTONE, Sonia et Claire LOCAL, « Measurement matters: Difficulties in defining and measuring children's television viewing in a changing media landscape », *Media International Australia*, vol. 163, n° 1, 2017, p. 67-76.
- MARRES, Noortje, « The redistribution of methods: On intervention in digital social research, broadly conceived », *The Sociological Review*, vol. 60, 2012, p. 139-165.
- MILLERAND, Florence, Christine THOËR et Caroline VRIGNAUD, « Regarder des séries en ligne: les formes de l'attachement chez de jeunes adultes québécois », dans Jean CHATEAUVERT et Gilles DELAUDAUD (dir.), *D'un écran à l'autre, les mutations du spectateur*, L'Harmattan, 2016, p. 557-571.
- MORLEY, David, *Family Television: Cultural Power and Domestic Leisure*, Comedia, Routledge, 1986.
- ORGAD, Shani, « From online to offline and back: Moving from online to offline relationships with research informants », dans Christine HINE (dir.), *Virtual Methods: Issues in Social Research on the Internet*, Berg, Oxford, 2005, p. 51-66.
- PASQUIER, Dominique, *L'Internet des familles modestes. Enquête dans la France rurale*, Presses des Mines, 2018.
- PASQUIER, Dominique, *De la culture des sentiments*, Éditions de la Maison des sciences de l'homme, 1999.
- PASQUIER, Dominique, Valérie BEAUDOUIN et Tomas LEGON, « *Moi, je lui donne 5/5* ». *Paradoxes de la critique amateur en ligne*, Presse des Mines, 2014.
- PROULX, Serge et Marie-France LABERGE, « Vie quotidienne, culture télé et construction de l'identité familiale », *Réseaux*, vol. 13, n° 70, 1995, p. 121-140.
- ROGERS, Richard, *Digital Methods*, The MIT Press, 2013.
- SILVERSTONE, Roger, Eric HIRSCH et David MORLEY, « Listening to a long conversation: An ethnographic approach to the study of information and communication technologies in the home », *Cultural Studies*, vol. 5, n° 2, 1991, p. 204-227.

- SNEE, Helene, Christine HINE, Yvette MOREY, Steven ROBERTS et Hayley WATSON (dir.), *Digital Methods for Social Science. An Interdisciplinary Guide to Research Innovation*, Palgrave Macmillan, 2016.
- SPIES, Virginie, « Twitter, l'avenir de la télévision ? », dans Jean CHATEAUVERT et Gilles DELAUDAUD (dir.), *D'un écran à l'autre, les mutations du spectateur*, L'Harmattan, 2016, p. 489-500.
- THOËR, Christine *et al.*, « Sur le Web, je regarde des vidéos, des séries et des émissions : Catégorisation des contenus de divertissement visionnés en ligne par les jeunes de 12 à 25 ans », *Comunicazioni Sociali*, n° 2, 2015, p. 191-207.
- VICENTE, Paula et Elizabeth REIS, « The "frequency divide" : Implications for Internet-based surveys », *Quality & Quantity*, vol. 47, n° 6, 2013, p. 3051-3064.

CHAPITRE 8

Les usages numériques adolescents : vers une ethnographie multisite de la « chambre numérique »

Nina Duque

L'adolescence est un processus transitionnel qui s'effectue, entre autres, par l'appropriation d'un espace intime, ici la chambre, à partir duquel les jeunes entrent en relation avec eux-mêmes et avec le monde (Lincoln, 2012). La chambre est leur premier espace, leur lieu d'appartenance, leur chez-soi. C'est au sein de cet endroit que les jeunes apprennent à articuler et à représenter leur vie sociale et culturelle (Glevarec, 2010). C'est là qu'ils vivent leurs premières expériences, créant et cimentant leur identité en y transposant désirs, goûts et aspirations. La chambre est un espace névralgique symbolisant le trajet de la découverte de soi. La culture de la chambre fait partie intégrante de ce trajet au cours duquel les jeunes s'approprient « leur » espace à travers l'affichage d'une identité, de passions et d'intérêts. Tel un laboratoire, la chambre devient un lieu d'expérimentations, de découvertes et de changements. Cette appropriation constitue la manifestation principale de l'exploration identitaire qui est à la fois individuelle et sociale (Pasquier, 2005; De Singly, 2006). De nos jours, les pratiques adolescentes se renouvellent et leurs modes communicationnels s'appuient aussi sur des dispositifs et des réseaux numériques (boyd, 2014). Ces mutations modifient non seulement les questionnements de recherche en culture de la chambre, mais également les méthodes de recherche qui s'y rattachent (Livingstone, 2007; Lincoln et Robards, 2014; Woodman et Bennett, 2015).

Il y a plusieurs enjeux méthodologiques liés à l'étude des usages que font les jeunes des technologies et médias numériques dans le cadre de la culture de la chambre, en particulier en ce qui concerne les approches méthodologiques s'inspirant des grands principes de la méthode ethnographique. Nous nous intéressons ici aux méthodes de recherche qui explorent les usages numériques adolescents, en particulier celles qui se rapportent à l'ethnographie du virtuel et qui permettent d'envisager la culture de la chambre comme moyen d'interpréter les pratiques numériques des jeunes. Nous examinons aussi ces méthodes à travers le potentiel des approches participatives et multisite en études de la culture de la chambre numérique.

L'ethnographie de la chambre

Dans sa forme la plus caractéristique, [l'ethnographie] repose sur la participation de l'ethnographe, ouvertement ou secrètement, à la vie quotidienne des personnes pendant une période prolongée, surveillant ce qui se passe, écoutant ce qui se dit, posant des questions en fait, rassemblant toutes les données disponibles pour apporter un éclairage sur les questions de recherche (traduction libre de Hammersley et Atkinson, 1995, p. 1)

La recherche qualitative mise en œuvre dans les travaux sur la jeunesse nous permet de saisir la richesse de l'expérience adolescente. Plus spécifiquement, l'ethnographie occupe une place privilégiée dans l'étude de la culture de la chambre. Aujourd'hui, plus de quarante ans après sa publication, l'article « *Girls and Subcultures* » de McRobbie et Garber (1976), considéré comme l'un des textes fondateurs sur la culture de la chambre, demeure une référence des plus connues et citées. Il existe d'ailleurs une littérature abondante qui considère l'ethnographie comme l'une des approches à privilégier pour explorer les contextes sociaux et culturels des jeunes (Goguel d'Allondans et Lachance, 2014). Issue de la tradition anthropologique, l'ethnographie consiste en l'exploration des pratiques dans leur réalité ordinaire, celle de tous les jours. Comme l'explique Lincoln, « [p]our l'ethnographe, c'est l'ordinaire qui est extraordinaire, et ce qui est extraordinaire n'apparaît que par l'observation attentive et approfondie du terrain » (traduction libre de Lincoln, 2012, p. 42). L'approche ethnographique demande que l'on s'immerge dans le groupe étudié et que l'on observe les activités et les interactions dans la vie normale des individus

sur une longue période de temps. De telles recherches peuvent ainsi enrichir les connaissances sur les usages et pratiques des jeunes tout en remettant en question certaines idées préconçues sur leur comportement (Cohen, 1993). Selon Bennett (2000), dont la recherche ethnographique sur les jeunes, la musique et les styles a permis d'acquérir de nombreuses connaissances sur l'identité et la socialisation adolescente, l'ethnographie est essentielle afin d'appréhender avec justesse la vie quotidienne des jeunes.

L'étude de la culture de la chambre conjugue des approches ethnographiques variées, notamment visuelles, multisensorielles et d'observation. Elle se réalise à partir d'outils aussi divers que des entretiens, des groupes de discussion, des carnets d'activités, des photographies de chambres et des cartographies de l'emplacement et de la disposition des objets dans la chambre. La recherche sur la chambre adolescente propose de ce fait une approche centrée sur l'observation des espaces et sur l'analyse du sens qu'ils ont pour les jeunes. Les groupes de discussion permettent aux jeunes de catégoriser leurs pratiques selon des critères qui leur sont propres et de réfléchir collectivement sur leurs pratiques. Ces données proviennent d'un dialogue entre jeunes plutôt que d'une discussion avec l'ethnographe. Les carnets d'activités, qu'ils soient écrits, sonores ou visuels, sont une riche source d'informations, un complément aux entretiens permettant aux jeunes de noter les petits moments du quotidien qui s'oublie facilement en entrevue. Les carnets permettent également d'accéder à des pratiques plus intimes, des pratiques plus difficiles à observer *in situ*. Enfin, les chambres adolescentes sont aussi des espaces de pratiques esthétiques quotidiennes. Photographier, filmer et cartographier la chambre permet d'explorer en détail le rôle de celle-ci dans la vie des jeunes puisqu'il s'agit d'un espace où l'identité s'exprime et se met en scène surtout par le visuel.

La recherche participative

Ce fait – que nos données ne sont en réalité que notre propre interprétation des perceptions d'autrui – est voilé puisque la majorité de ce que nous voulons comprendre, que ce soit un événement particulier, un rituel, une coutume, une idée, peu importe, est prise comme une information contextuelle avant même que le sujet soit directement exploré (traduction libre de Geertz, 1973, p. 9).

Bien que l'ethnographie soit toute désignée pour l'étude de la culture de la chambre, elle n'est pas exempte de certaines remises en question. Selon Hodkinson (2017), de plus en plus de spécialistes menant des recherches sur les pratiques numériques adolescentes s'inquiètent de sa mise en œuvre. Lincoln (2012), pour sa part, suggère l'existence actuelle d'un « cul-de-sac » concernant l'interprétation et la représentation des données recueillies. Parmi les principales difficultés qu'elle mentionne, soulignons la partialité des récits, la subjectivité des points de vue et le rôle de l'ethnographe. Pour les critiques issues d'une pensée plus positiviste, les données recueillies ne reflètent pas la « vraie » réalité des jeunes. Ce corpus serait plutôt une version subjective du vécu construite par les personnes effectuant la recherche et y prenant part. D'ailleurs, c'est justement cette relation d'intimité entre les sujets de la recherche et les personnes qui la réalisent qui rend problématique l'ethnographie dans la chambre (Heath *et al.*, 2009). Certes, cette approche privilégie un engagement étroit dans les mondes adolescents, où les ethnographes établissent, par le partage de la vie quotidienne, des relations rapprochées, voire amicales et profondes. Nous avons éprouvé nous-mêmes – lors de nos recherches auprès des jeunes – un réel attachement (souvent réciproque).

À l'inverse des critiques positivistes, pour Bennett (2002), ce serait plutôt un manque d'engagement inhérent aux recherches sur les jeunes qui produirait des comptes rendus ethnographiques douteux, et ce, malgré le succès et l'influence que les études plus réalistes ont obtenus. En outre, un nombre grandissant d'études ethnographiques en culture de la chambre s'inscrivent dans une vision constructiviste et explorent des processus de recherche plus participatifs. On considère dès lors la recherche comme un processus situé, collaboratif et réflexif reconnaissant l'impossibilité de toute objectivité absolue. À juste titre, pour Mazzarella (2010), une démarche ethnographique dans la chambre adolescente se doit d'être souple et flexible. Et même si les données recueillies sont coproduites, le corpus issu de cette collecte est tout aussi juste et valide. Il suffit, dans le processus de recherche, de minimiser l'impact de l'ethnographe sur les données tout en s'efforçant – à l'intérieur d'un groupe ou d'un contexte – d'explorer et de saisir les comportements et le sens accordés aux actions des jeunes. Cette optique se rapproche de la pensée de Geertz (1973), qui avançait, il y a de cela près d'un demi-siècle, que pour bien comprendre la signification des actions des gens (les jeunes, dans notre cas), l'ethnographe

se doit d'être attentif aux individus et à leurs choix. Il s'agit de les suivre dans leurs pratiques en décrivant de manière fine et détaillée leurs expériences – toujours dans leur environnement – tout en mettant en relief l'ancrage de ces usages et de ces pratiques dans la logique de production des sujets. Il est important que les descriptions faites par l'ethnographe aient du sens pour les sujets. Pour ce faire, ceux-ci doivent pouvoir se reconnaître et se voir dans ces portraits. Enfin, toujours selon Geertz, nous devons être conscients du fait que nos analyses sont, à la base, notre construction de la réalité des autres.

Une réponse fort intéressante à ces critiques consiste à placer les jeunes de manière qu'ils puissent « raconter leur propre histoire » en leur accordant un rôle clé dans la construction des données recueillies et dans leur analyse. Cette perspective, inspirée des recherches en éducation orientées vers la participation active des jeunes, leur offre la possibilité d'être des agents actifs plutôt que passifs dans l'élaboration et la mise en œuvre des études sur leurs pratiques. Ce qui caractérise cette approche « participative » ne tient pas tant aux méthodes utilisées qu'au degré d'implication des jeunes dans l'ensemble du processus de recherche. Par exemple, nous avons choisi d'élaborer notre projet de recherche doctorale de façon coopérative. Nous prévoyons de négocier les objectifs de notre recherche avec nos jeunes participants, et les questions et les analyses seront collaboratives. En invitant les jeunes dans les coulisses de notre projet, nous pourrions déterminer, nous l'espérons, les rapports et les processus vécus par les jeunes et dans lesquels ils s'inscrivent tout en les engageant dans la découverte de ces processus. En ce sens, une recherche participative devient une double source de données permettant des constatations et des observations plus poussées.

Les chambres numériques

À certains égards, les publics médiatisés sont comparables aux publics non médiatisés que la majorité connaît bien : on les voit dans les parcs, les centres commerciaux, les parcs de stationnement, les cafés, etc. Les adolescents s'y présentent afin de rejoindre leurs amis. D'autres personnes peuvent être présentes et intervenir dans la conversation si elles sont considérées comme intéressantes, ou ignorées si elles ne le sont pas. (traduction libre de boyd, 2014, p. 2).

Depuis l'adoption de la communication médiatisée par ordinateur à domicile, des universitaires de traditions et courants de recherche divers s'attellent à comprendre les pratiques, les usages et les espaces en ligne propres aux publics adolescents. Le recours à la métaphore de la chambre comme moyen de saisir les espaces et les usages adolescents en ligne s'amorce dans les années 1990, en particulier avec la popularité croissante d'Internet à domicile et l'émergence de sites Web personnels. Les premières études, s'inscrivant dans le courant des études culturelles, font un lien entre la présentation en ligne d'intérêts personnels, de goûts, de préférences ou d'amitiés et la façon dont les jeunes affichent ces mêmes marqueurs identitaires aux murs de leur chambre (Steele et Brown, 1995; Larson, 1995). Dans son étude sur les usages des médias dans les chambres, Arnett (1995) se penche sur les conséquences d'une vie de famille de plus en plus privatisée et de l'accès sans précédent aux médias. Selon lui, les pratiques médiatiques « modernes » mèneraient à un processus « d'auto-socialisation » chez les jeunes. Pour Livingstone (2007), c'est plus précisément la crainte du « monde extérieur » qui encourage les parents à se procurer des technologies médiatiques et à les introduire dans les espaces domestiques adolescents comme la chambre.

S'ensuivent au fil des années des études sur les espaces connectés adolescents qui abordent des sujets aussi divers que l'expression et l'exploration des interactions au sein de cercles amicaux (Itō *et al.*, 2010) ou la performance identitaire en ligne (Bennett et Kahn-Harris, 2004). Ces dernières s'intéressent, entre autres, aux pages Web, à la construction du sol, à la création d'espaces participatifs ou aux usages sociaux. Certaines études se sont fortement inspirées de l'analogie de la chambre en soutenant que les médias numériques peuvent être considérés comme une série d'espaces connectés utilisés (et perçus) par les jeunes de la même manière que leur chambre (Hodkinson et Lincoln, 2008; Robards et Bennett, 2011). Par exemple, l'étude de Pearson (2009) sur les performances identitaires en ligne révèle la façon dont les échanges en ligne s'apparentent aux conversations entre amis dans une chambre. La forte popularité aujourd'hui des systèmes de messagerie comme Snapchat ou WhatsApp soutient cette idée d'échange et de partage social entre pairs. Pour Tiidenberg (2015), qui s'intéresse aux conflits en ligne, les jeunes considèrent les plateformes telles que celle de Tumblr comme des espaces sûrs dans lesquels il est possible d'avoir des interactions sans danger avec un ensemble d'amis

choisis. Lincoln et Robards (2014), pour leur part, explorent les affichages identitaires sur Facebook et affirment que de tels sites participent à la construction identitaire à long terme. Hodkinson (2017) abonde dans leur sens en suggérant que les réseaux sociaux constituent des territoires personnels permettant aux jeunes d'explorer et de construire leur identité. Ce sentiment de sécurité et d'appartenance pourrait expliquer le fait que les jeunes se sentent « chez eux » sur les réseaux sociaux, tout comme dans leur chambre.

Il se produit donc, au sein des études sur la culture de la chambre, un renouvellement de perspective où la chambre n'est plus envisagée exclusivement dans sa dimension physique ; de nouveaux espaces numériques, connectés et accessibles à tout moment s'y intègrent. Dans ces « chambres numériques », les jeunes explorent tout autant des modèles identitaires, construisent leur soi et socialisent entre pairs (Duque, 2017). Ces transformations dans les espaces vécus et habités par les jeunes nous forcent à repenser nos méthodes d'enquête. Si la chambre n'est plus exclusivement fixée dans l'enceinte de la maison, ses nouvelles frontières poreuses exigent le développement de méthodes « hybrides » qui permettront une analyse des lieux adolescents qui existent de nos jours autant en ligne que hors-ligne.

L'ethnographie virtuelle et la chambre numérique

L'ethnographie a été intégrée à la culture universitaire en tant que moyen approprié d'explorer la manière dont les gens donnent un sens aux possibilités offertes par Internet (traduction libre de Hine, 2000, p. 260).

Les spécialistes de l'étude de la chambre numérique s'inspirent fortement des ethnographies virtuelles. Ces méthodes, spécifiquement adaptées aux mondes numériques, se sont développées en réaction aux limites perçues de l'ethnographie traditionnelle appliquée aux mondes numériques et offrent ainsi une occasion de renouveler les façons de faire des recherches. Selon boyd (2014), les espaces en ligne ont des caractéristiques particulières qui nécessitent des approches adaptées. Parmi les particularités des mondes numériques qui ont motivé ces changements se trouvent l'anonymat, la contraction de l'espace et du temps, la présence de publics invisibles ou la persistance des mondes interactifs. Ces affordances s'accompagnent de trois éléments sous-jacents : (1) un auditoire invisible (la

possibilité de voir sans être vu), (2) des contextes changeants et l'impossibilité pour l'utilisateur et l'usagère d'adapter son comportement adéquatement, et, (3) le brouillage des frontières privé/public. Pour plusieurs, dont Hine (2000) et Kozinets (2015), les affordances des technologies numériques contribuent à développer une dynamique (et une culture) propre qui nécessite des méthodes et outils de recherches taillés sur mesure.

L'ethnographie virtuelle se distingue de l'ethnographie classique en ce qu'elle s'intéresse spécifiquement à ce qui est produit en ligne par un individu. On pense à l'étude des communautés sur les réseaux numériques tels que les groupes et les forums de discussion, les sites Internet ou les mondes virtuels, les conversations suscitées par les blogues ou les échanges sur des forums de discussion (Hodkinson et Lincoln, 2008) ; à l'exploration de la mise en scène sur des pages personnelles ou publiques (boyd, 2014) ; à l'analyse de traces numériques et à l'observation en ligne (Marwick et boyd, 2011). On insiste sur la notion de participation, de partage et d'apprentissage de codes, de procédures, de lois, de règles et de routines. Bennett (2004) note que l'ethnographie virtuelle fournit ainsi en abondance des occasions de découvrir de nouveaux espaces et de nouvelles cultures propres à la vie adolescente.

Cette approche centrée sur les pratiques en ligne n'est pas sans susciter elle aussi quelques critiques. S'il peut s'agir d'une révolution, notamment pour Markham et Baym (2008), d'autres n'y voient aucune différence avec l'ethnographie traditionnelle, ni même la nécessité d'une approche distincte (Beaulieu, 2004). Comme l'énonce Pastinelli (2011, p. 40),

Presque toutes ces caractéristiques qu'on attribue à l'espace électronique peuvent également être attribuées à certains contextes hors-ligne et on peut aussi, inversement, trouver dans le vaste répertoire des pratiques sociales supposant l'usage des communications électroniques des contre-exemples pour chacune de ces caractéristiques.

Enfin, si les recherches explorant la culture de chambre numérique et privilégiant des méthodes ethnographiques virtuelles s'intéressent à juste titre à ce que les jeunes accomplissent en ligne, elles le font parfois à l'exclusion des pratiques ou des contextes hors-ligne. On analyse les traces numériques (textes, images, photos, vidéos, profils et identités numériques) laissées par les jeunes lors de leurs pratiques en ligne, sans méthodiquement les resituer dans l'ensemble des pratiques qui prennent place dans la chambre. Or, nous croyons qu'il faut faire attention de ne

pas réduire la chambre numérique à la seule production de traces ou aux seuls mondes connectés. Comme le souligne avec justesse Driscoll, « [il faut y] inclure la vie de tous les jours, car c'est ce "quotidien" [...] qui rend la participation en ligne possible et significative » (traduction libre, 2002, p. 17).

Vers une ethnographie multisite de la chambre numérique

[Au] sein d'un imaginaire de recherche multisite, la recherche et la description des liens et des relations entre des sites jusque-là jugés incomparables constituent les outils de l'ethnographe pour formuler ses arguments et construire ses propres contextes de signification (traduction libre de Marcus, 1998, p. 14)

Comme nous l'avons vu, les jeunes ne se cantonnent plus à une chambre en particulier, mais traversent et habitent une multitude d'espaces (numériques et physiques, publics et privés). Chez certains, les pratiques s'effectuent sur une seule plateforme (par exemple, YouTube) dans un seul lieu (la chambre), alors que pour d'autres, elles se font à travers divers espaces en ligne (Facebook, Instagram, Twitter) et hors-ligne (la chambre, l'école, les parcs, les transports en commun). De la sorte, l'étude de la culture de la chambre à l'ère numérique pourrait s'appuyer sur une ethnographie multisituée adaptée à la fois aux espaces hors-ligne et en ligne et combinant des méthodes traditionnelles et virtuelles.

L'ethnographie multisite, telle que la décrit Marcus (1995), propose justement un riche cadre de référence nous permettant d'analyser les mondes adolescents contemporains. Il s'agit de découvrir de nouvelles connexions, associations et interactions s'exprimant sur un terrain configuré autrement. Comme le souligne Hine (2007), l'approche multisite permet de représenter fidèlement le vécu non pas dans des lieux distincts, mais plutôt à travers les connexions qu'elles produisent. Présentant un intérêt certain pour notre recherche, l'ethnographie multisite permet de poursuivre en parallèle un travail de terrain ethnographique dans plus d'un lieu sans compromettre la représentativité ou la densité des données recueillies. Si la culture de la chambre « traditionnelle » implique une recherche liée à un espace déterminé (la chambre), une approche multisituée nous permettra de construire une ethnographie capable de tracer et d'établir des liens entre des pratiques et des usages adolescents qui se déploient dans une pluralité de mondes. Une telle approche permet de

quitter le cadre restrictif d'un terrain unique – limité à la chambre ou à un réseau numérique – et de suivre les jeunes dans leurs déplacements d'un endroit à l'autre et dans les différents médias qu'ils privilégient.

Pour brosser un portrait plus complet de la culture de la chambre dans l'ensemble des mondes où elle est vécue aujourd'hui, il faut la resituer dans l'ensemble des pratiques sociales adolescentes, autrement dit suivre les jeunes dans tous les environnements qu'ils habitent (les chambres physiques et numériques). Tout en s'intéressant aux dispositifs et aux usages numériques, il faut s'intéresser avant tout à l'expérience vécue par les jeunes et au sens qu'ils attribuent à leurs pratiques, relations et interactions dans l'unité qu'est leur vie quotidienne.

Rendre compte des pratiques adolescentes en ligne et de la relation de ces jeunes à leur quotidien hors-ligne pose un défi de taille. Nous croyons que l'exploration des approches ethnographiques traditionnelles, virtuelles et multisite peut nous aider à imaginer des méthodes inédites de recherche en culture de la chambre.

S'il y a une chambre adolescente de nos jours, elle se déploie dans plusieurs espaces dont l'ethnographe doit tenir compte. Ainsi, l'étude de la chambre doit prendre en compte l'aspect multisitué des pratiques adolescentes. Ce faisant, nous pourrions éviter une segmentation des lieux et des mondes adolescents et, surtout, ne pas confiner leurs pratiques dans des carcans mutuellement exclusifs. Une approche adéquate devra tenir compte de la complexité de l'étude de l'intersection, dans le temps et l'espace, des chambres numériques et physiques. Finalement, pour témoigner le plus fidèlement possible du vécu des jeunes et comprendre leurs expériences, il importe d'adopter une méthodologie participative et adaptée aux réalités adolescentes qui, seule, peut permettre de rendre compte de la « nouvelle » chambre adolescente.

Références

- ARNETT, Jeffrey J., « Adolescents' uses of media for self-socialization », *Journal of Youth and Adolescence*, vol. 24, n° 5, 1995, p. 519-533.
- BEAULIEU, Catherine, « Intercultural study of personal space: A case study », *Journal of Applied Social Psychology*, vol. 34, n° 4, 2004, p. 794-805.
- BENNETT, Andy, « Researching youth culture and popular music: A methodological critique », *British Journal of Sociology*, vol. 53, n° 3, 2002, p. 451-466.
- BENNETT, Andy, *Popular Music and Youth Culture: Music, Identity and Place*, Palgrave Macmillan, 2000.

- BENNETT, Andy et Keith KAHN-HARRIS, « Virtual subculture? Youth, identity and the Internet », dans *After Subculture: Critical Studies in Contemporary Youth Culture*, Palgrave Macmillan, 2004, p. 162-172.
- BOYD, danah, *It's Complicated: The Social Lives of Networked Teens*, Yale University Press, 2014.
- COHEN, Sara, « Ethnography and popular music studies », *Popular Music*, vol. 12, n° 2, 1993, p. 123-128.
- DE SINGLY, François, « Le sens de la chambre personnelle pendant la seconde modernité. Le cas de l'adonnaissance », dans Phuong Mai HUYNH (dir.), *Habitat et vie urbaine: changements dans les modes de vie*, Puca, 2006, p. 32-44.
- DRISCOLL, Catherine, *Girls: Feminine Adolescence in Popular Culture and Cultural Theory*, Columbia University Press, 2002.
- DUQUE, Nina, « La reconsidération de la "culture de chambre" : une appropriation intime des espaces numériques par les adolescentes », dans Maude BONENFANT, Fabien DUMAIS, Gabrielle TRÉPANIÉRIE-JOBIN (dir.), *Les pratiques transformatrices des espaces numériques*, Cahiers du Gerse, Presses de l'Université du Québec, 2017.
- GEERTZ, Clifford, « Thick description: Toward an interpretive theory of culture », dans *The Interpretation of Cultures*, Basic Books, 1973, p. 3-33.
- GLEVAREC, Hervé, *La culture de la chambre: Préadolescence et culture contemporaine dans l'espace familial*, La Documentation Française, 2010.
- GOGUEL D'ALLONDANS, Thierry et Jocelyn LACHANCE, *Étudier les ados*, Collections politiques et interventions sociales, Presses de l'EHESP, 2014.
- HAMMERSLEY, Martyn et Paul ATKINSON, *Ethnography: Principles in Practice*, Psychology Press, 1995.
- HEATH, Sue *et al.*, *Researching Young People's Lives*, SAGE Publications, 2009.
- HINE, Christine, « Multi-sited ethnography as a middle range methodology for contemporary STS », *Science, Technology & Human Values*, vol. 32, n° 6, 2007, p. 652-671.
- HINE, Christine, *Virtual Ethnography*, SAGE Publications, 2000.
- HODKINSON, Paul, « Bedrooms and beyond: Youth, identity and privacy on social network sites », *New Media & Society*, vol. 19, n° 2, 2017, p. 272-288.
- HODKINSON, Paul et Siân LINCOLN, « Online journals as virtual bedrooms? Young people, identity and personal space », *YOUNG: The Nordic Journal of Youth Research*, vol. 16, n° 4, 2008, p. 625-650.
- ITŌ, Mizuko *et al.*, *Hanging Out, Messing Around, and Geeking Out: Kids Living and Learning with New Media*, The MIT Press, 2010.
- KOZINETS, Robert V., « Netnography », dans *The International Encyclopedia of Digital Communication and Society*, 2015.
- LARSON, Reed, « Secrets in the bedroom: Adolescents' private use of media », *Journal of Youth and Adolescence*, vol. 24, n° 5, 1995, p. 535-550.
- LINCOLN, Siân, *Youth Culture and Private Space*, Palgrave Macmillan, 2012.
- LINCOLN, Siân et Brady ROBARDS, « Being strategic and taking control: Bedrooms, social network sites and the narratives of growing up », *New Media & Society*, vol. 18, n° 6, 2014, p. 927-943.

- LIVINGSTONE, Sonia, «From family television to bedroom culture: Young people's media at home», dans Eoin DEVEREUX (dir.), *Media Studies: Key Issues and Debates*, SAGE Publications, 2007, p. 302-321.
- MARCUS, George E., «Introduction», dans *Ethnography through Thick & Thin*, Princeton University Press, 1998, p. 3-30.
- MARCUS, George E., «Ethnography in/of the world system: The emergence of multi-sited ethnography», *Annual Review of Anthropology*, vol. 24, 1995, p. 95-117.
- MARKHAM, Annette N. et Nancy K. BAYM, *Internet Inquiry: Conversations about Method*, SAGE Publications, 2008.
- MARWICK, Alice E. et danah BOYD, «I tweet honestly, I tweet passionately: Twitter users, context collapse, and the imagined audience», *New Media & Society*, vol. 13, n° 1, 2011, p. 114-133.
- MAZZARELLA, Sharon R., *Girl Wide Web 2.0*, Peter Lang, 2010.
- MCRROBBIE, Angela et Jenny GARBER, «Girls and subcultures», dans Stuart HALL et Tony JEFFERSON (dir.), *Resistance Through Rituals: Youth Subcultures in Post-War Britain*, Routledge, 1976, p. 209-223.
- MILES, Matthew B. et A. Michael HUBERMAN, *Qualitative Data Analysis (2nd edition)*, SAGE Publications, 1994.
- PASQUIER, Dominique, *Cultures lycéennes: la tyrannie de la majorité*, Autrement, 2005.
- PASTINELLI, Madeleine, «Pour en finir avec l'ethnographie du virtuel! Des enjeux méthodologiques de l'enquête de terrain en ligne», *Anthropologie et Sociétés*, vol. 35, n°s 1-2, 2011, p. 35-52.
- PEARSON, Erika, «All the World Wide Web's a stage: The performance of identity in online social networks», *First Monday*, vol. 14, n° 3, 2009.
- ROBARDS, Brady et Andy BENNETT, «MyTribe: Manifestations of belonging on social networking sites», *Sociology*, vol. 45, n° 2, 2011, p. 303-317.
- STEELE, Jeanne R. et Jane D. BROWN, «Adolescent room culture: Studying media in the context of everyday life», *Journal of Youth and Adolescence*, vol. 24, n° 5, 1995, p. 551-576.
- TIIDENBERG, Katrin, «Boundaries and conflict in a NSFW community on Tumblr: The meanings and uses of selfies», *New Media & Society*, vol. 18, n° 8, 2015, p. 1563-1578.
- WOODMAN, Dan et Andy BENNETT, *Youth Cultures, Transitions, and Generations. Bridging the Gap in Youth Research*, Palgrave Macmillan, 2015.

CHAPITRE 9

L'observation des représentations adolescentes de l'intimité : enquête sur YouTube

Claire Balleys

YouTube est une plateforme de partage de contenus qui permet de publier, d'évaluer et de commenter des vidéos. Plébiscitée notamment par les jeunes, elle a donné naissance à des pratiques inédites de présentation de soi et d'affiliation à des communautés en ligne¹, centrées sur un besoin d'appartenance et d'identification (Rotman et Preece, 2010 ; Raun, 2012 ; Lange, 2014 ; Balleys, 2017a). Les vidéos postées par des jeunes sont considérées comme un nouveau genre de « supports d'expression personnelle » (Lange, 2014) à vocation participative puisqu'elles sont destinées à être vues et évaluées par un public large, si possible (Caron, 2014 ; Balleys, 2017b). Leur signification est coconstruite par les personnes qui les créent et leur public (Lange, 2014, p. 27) qui se retrouvent autour d'intérêts, de goûts, d'expériences ou de sentiments d'appartenance communs (Rotman et Preece, 2010).

Dans ce chapitre, nous proposons de dérouler le fil réflexif qui a guidé la constitution puis le travail d'analyse d'un terrain ethnographique sur la mise en scène de soi et de son intimité par des jeunes, garçons et filles,

1. Par communautés en ligne, nous entendons « un groupe ou plusieurs sous-groupes de personnes, réunies par un intérêt partagé, utilisant une plateforme médiatisée, pour interagir et créer des contenus générés par les usagers, qui sont accessibles à tous les membres de la communauté, afin de cultiver une culture commune et d'adhérer à des normes communes » (traduction libre de Rotman et Preece, 2010, p. 320).

sur YouTube. L'objectif est de revenir sur le processus d'élaboration de la problématique et sur la manière dont il s'est articulé avec la construction du terrain d'enquête : une immersion ethnographique de dix-huit mois suivie de la stabilisation d'un corpus de 80 vidéos francophones réalisées par des garçons et des filles entre 14 et 18 ans, vivant en France, en Belgique, en Suisse romande et au Québec.

Comme l'indiquent Beaud et Weber, le travail ethnographique « se présente comme une tension entre familiarité et étrangeté » (2003, p. 47). Le défi est de « mettre à jour la complexité des pratiques les plus ordinaires des enquêtés » (*ibid.*, p. 9), c'est-à-dire de déconstruire ce qui semble aller de soi, mais aussi de « comprendre les discours étranges » (Becker, 2002, p. 239), soit de découvrir le sens de ce qui relève du langage propre à la population ou à la communauté étudiée, qui n'est d'ordinaire accessible qu'à leurs membres. Le principe ethnographique repose sur une immersion dans une culture ou dans un « monde » (Berry, 2012) qu'il s'agit de pénétrer et de comprendre, en révélant ce sur quoi reposent ses logiques sociales et articulations internes. Beaud et Weber rappellent également que « faire du terrain revient à rendre justice à des pratiques ignorées, mal comprises ou méprisées », voire à les réhabiliter (2003, p. 9). Cette affirmation rejoint les raisons qui nous ont conduites à travailler sur les pratiques adolescentes de présentation de soi sur YouTube. En effet, ce que font les adolescents et les adolescentes sur les médias sociaux est *a priori* considéré par les adultes comme étrange, voire vain ou dangereux (Balley, 2015). Plus qu'une volonté de réhabilitation des pratiques de sociabilité et de mise en scène de soi en ligne, l'objectif de cette recherche a été de rendre visibles et de faire connaître des usages qui sont souvent ignorés par les adultes – individus ou institutions – en les situant dans leur contexte, à savoir les processus contemporains de socialisation adolescente.

Si une multitude d'informations sont disponibles sur YouTube, toutes ne permettent pas d'offrir des données sociologiques pertinentes. Le cœur de notre propos est de proposer une réflexion sur le point de rencontre entre la problématique de recherche et la focale ethnographique sur YouTube. Autrement dit, qu'est-ce que ce terrain d'investigation a à offrir à la recherche et que va-t-il dissimuler ? Que va-t-on pouvoir en extraire et pour produire quel type de connaissance ? Les réponses à ces questions ne peuvent être fournies au début de l'enquête, car si « l'objet fait le ter-

rain», il est inévitable, pour l'ethnographie, que le terrain fasse l'objet (Beaud et Weber, 2003, p. 50). Il en résulte que le travail empirique fait émerger les questions de recherche au fil de l'investigation.

Ce constat est encore plus valable en ce qui concerne le travail ethnographique en ligne, qui est plus difficile à circonscrire que l'ethnographie en présentiel (Hine, 2008). Par exemple, en choisissant d'établir son terrain dans une cour de récréation ou un club de tennis, l'ethnographe se trouve face à un espace-temps prédéfini, ce qui n'est pas le cas lors d'une enquête sur Internet. À la nécessité de fixer le contexte spatio-temporel de son investigation s'ajoute la difficulté à cibler précisément les contenus. Pour ce faire, nous proposons d'ancrer solidement l'observation ethnographique en ligne dans un cadre conceptuel maîtrisé. En effet, si la plasticité de l'ethnographie en ligne implique « une dimension de bricolage méthodologique et de créativité » (Jouët et Le Caroff, 2013, p. 148), la boîte à bricolage doit contenir de bons outils pour que le résultat soit scientifiquement significatif.

L'intimité au cœur d'enjeux de socialisation à l'adolescence

Cette enquête ethnographique sur YouTube a été élaborée et construite à partir du constat suivant : la notion d'intimité est au cœur d'importants enjeux identitaires au moment de l'adolescence (Balleys, 2015 ; Balleys et Coll, 2015). Premièrement, les liens forts tissés entre pairs constituent une marque de prestige à l'adolescence, car ils sont symboles d'autonomisation et, donc, de maturité (Metton-Gayon, 2009 ; Balleys, 2015). Avoir un ami proche ou une première partenaire amoureuse permet de faire la démonstration d'une capacité à recevoir de l'affection et de la reconnaissance de la part d'« autrui significatif » choisis (et conquis) et non plus donnés (Berger et Luckman, 2008). La socialisation adolescente passe ainsi par la gestion et la valorisation de liens intimes, c'est-à-dire personnalisés, privilégiés et exclusifs entre pairs (Balleys, 2015). Dans ce contexte, la personne qui est désignée par ses pairs comme étant « sans ami » représente la figure du dépourvu social (Castel, 2006), incapable d'intégration. L'insulte la désigne à la fois comme une intruse et une incompétente (Balleys, 2016a). Si ces « logiques de l'exclusion » (Elias et Scotson, 1994) ne sont pas corrélées à l'avènement des médias sociaux, ces derniers leur ont néanmoins offert de nouvelles plateformes de déploiement et de mise

en visibilité. Le travail de valorisation de soi effectué par les jeunes en ligne passe par la valorisation d'une forme d'intimité, en particulier relationnelle: «Le degré "d'authenticité" de son intimité que l'on parvient à faire valider publiquement devient le critère de reconnaissance de sa valeur propre» (Balleys et Coll, 2015, p. 14).

Parallèlement à la production de ces résultats, nous avons commencé à nous intéresser à la plateforme YouTube, sur laquelle chaque jour des milliers de jeunes publient des vidéos dans lesquelles ils se mettent en scène face à la caméra depuis l'intimité de leur chambre (Balleys, 2016b). Ces productions médiatiques correspondent à des formats très différents de ceux qui caractérisent les échanges ayant lieu entre amis ou entre camarades d'école sur les médias sociaux, puisqu'il s'agit ici d'interagir avec un public non défini et dépassant le cercle des interconnaissances (Raun, 2012; Lange, 2014; Balleys, 2016b). Néanmoins, les premiers exercices d'immersion ethnographique ont révélé que l'un des formats dominants sur YouTube est un format de confession intime. Premièrement, les vidéos présentent des discours sur soi qui abordent des thématiques intimes – la puberté, la masturbation, les relations de couple, les ruptures amoureuses –, c'est-à-dire de révélation de ce qui, «de l'expérience et de l'identité de l'individu, n'est pas visible ou saisissable de l'extérieur par autrui, et donc qui appartient en propre à la subjectivité individuelle» (Latzko-Toth et Pastinelli, 2013, p. 156). Ce faisant, l'individu scénarise ce qui est de l'ordre du «caché» et de «l'au-delà de ce qu'on dit en public» (Glevarec, 2014, p. 37). Il y a donc un exercice consistant à publier l'intime observable sur YouTube, inscrit dans la lignée du mouvement opéré par les séries américaines au tournant des années 1990. Deuxièmement, la mise en scène des récits prend forme dans un contexte intimiste: le YouTubeur – autant que la YouTubeuse – s'exprime depuis sa chambre, s'adresse à son public face à la caméra, en le tutoyant comme s'il s'agissait d'une conversation entre proches.

À partir de ces premières observations a émergé la question suivante: quelles dimensions de l'intimité sont problématisées par les adolescents YouTubeurs et YouTubeuses? Les contenus médiatiques analysés, à savoir des vidéos publiées sur le site de partage, sont le fruit du travail d'appropriation d'un format culturel existant, popularisé à la télévision puis sur le Web. Il a été démontré que les processus de réception des produits culturels sont corrélés à des sentiments d'appartenance commune, ou au

contraire à des mécanismes d'exclusion (Dayan, 1992; Pasquier, 1999; 2005). Comme l'a énoncé Dayan, « être spectateur, c'est faire partie d'un "nous", mais ce "nous" se constitue en s'opposant à des "autres" et aux lectures que ces autres font ou qu'on leur prête » (1992, p. 14-15). Aimer un genre culturel ou un « style » musical est une manière de s'affilier à une communauté de réception qui partage ses goûts et ses opinions. Les pratiques de visionnement sur YouTube sont inscrites dans la même logique, d'autant qu'elles sont éminemment dialogiques et participatives (Lange, 2014; Burgess et Green, 2018). Leurs publics ont en effet la possibilité de manifester leur affiliation ou leur désaffiliation de manière très pragmatique, que ce soit par les mentions « j'aime » ou « je n'aime pas » associées à chaque vidéo, ou par la rédaction de commentaires (Beuscart et Mellet, 2015). Les jeunes sur YouTube sont en constante demande de réactions de la part de leur public, que ce soit par la souscription d'un abonnement à leur chaîne, l'expression de marques de soutien, d'avis ou de suggestions, par des sollicitations ritualisées à la participation : « si cette vidéo vous a plu, mettez-moi un petit pouce bleu, ça fait toujours plaisir » ; « n'oubliez pas de vous abonner à ma chaîne » ; « dites-moi dans les commentaires si vous voulez plus de vidéos de ce genre ».

En résumé, s'il est scientifiquement admis que YouTube est un espace participatif d'un point de vue social, identitaire, culturel et politique (Burgess et Green, 2018), le défi est de définir et d'explorer un terrain ethnographique qui donne accès aux contenus et aux significations de ces interactions. Comment l'intimité constitue-t-elle une ressource symbolique dans l'appropriation de cet espace participatif ?

La construction du terrain ethnographique sur YouTube

Pour répondre à ce défi, nous avons mené un travail de collecte puis d'analyse de données qui a été effectué en deux étapes successives et complémentaires entre les mois de février 2015 et de décembre 2016. Comme nous le détaillons, ce sont finalement deux terrains méthodologiquement distincts, quoique étroitement articulés, que nous avons conduits durant cette période.

On a commencé le premier terrain en délimitant le genre de vidéo pertinente pour affiner nos premières observations. Partant du présupposé que l'intimité est mobilisée sur YouTube comme une ressource symbolique

pour parler de soi en tant que membre d'une certaine communauté, elle devait par conséquent constituer l'une des composantes des vidéos intégrées à notre navigation sur la plateforme. Comme il a été mentionné, les phases exploratoires d'immersion ethnographique ont révélé que l'intimité est une thématique dominante dans les vidéos adolescentes, qu'elle soit subjective, relationnelle ou corporelle. S'il existe une grande diversité de catégories de vidéos sur YouTube, certaines deviennent rapidement populaires et constituent des canevas qui sont largement repris. Par exemple, les vidéos intitulées « cher corps », dans lesquelles on s'adresse à son corps de manière bienveillante (comme à une amie), sont un type de vidéo populaire chez les filles. Dans ce contexte, le rapport subjectif au corps est mobilisé pour créer un discours sur soi adressé à un public que l'on prend à témoin et dont on sollicite la participation, notamment par la publication de commentaires sur son propre rapport au corps. Le partage d'expériences intimes est au cœur du processus de mise en scène de soi sur YouTube, mais s'inscrit toujours dans un collectif qui dépasse l'individu et auquel les jeunes vidéastes font explicitement appel. Le procédé est donc très différent du rapport introspectif qui caractérise, par exemple, l'écriture d'un journal intime.

Afin de nous assurer d'avoir affaire à des contenus médiatiques pensés et construits pour être publiquement accessibles, nous avons pris la précaution de ne visionner que des vidéos qui, même si elles ne totalisent qu'une centaine de vues, ont pour vocation d'être diffusées le plus largement possible. Premièrement, nous nous sommes concentrées sur des vidéos qui ont été l'objet d'un travail d'écriture, de réalisation scénaristique et de post-production. Il se trouve que les jeunes font souvent référence au temps passé à préparer et à réaliser leurs vidéos, notamment pour souligner l'importance de leur implication dans la pratique. L'observation ethnographique s'est limitée aux vidéos dans lesquelles un travail de stylisation est explicite : le montage est visible, le décor comme le cadrage sont soignés, la luminosité, propre. Nous avons ainsi exclu toutes les vidéos filmées sur un mode spontané, ce qui est généralement énoncé explicitement et où la caméra est parfois allumée dans la pénombre de la chambre, voire du lit, et le débit de la voix est hésitant, mal assuré. Deuxièmement, les vidéos visionnées devaient comprendre une mention du désir de diffusion médiatique et de leur volonté d'acquérir de la popularité grâce à YouTube : « Partagez cette vidéo ! » ; « Dites-moi dans les commentaires si vous voulez d'autres vidéos de ce genre ! » ou ; « N'hésitez pas à partager la vidéo sur les réseaux sociaux ! ».

Tout au long du travail d'observation, nous avons conservé un strict rôle d'observatrice (Gold, 2003), c'est-à-dire invisible et non participant. Le seul acte d'engagement sur YouTube a été l'abonnement à 411 chaînes de garçons et de filles, âgés entre 14 et 18 ans. Nous avons évité les recherches thématiques par entrées de mots clés liés à la thématique de l'intimité, car un tel choix aurait limité les résultats aux propositions algorithmiques de YouTube. Nous avons préféré naviguer grâce aux liens insérés dans les commentaires de vidéos publiées par des vidéastes populaires. En effet, lorsqu'un YouTubeur ou une YouTubeuse célèbre publie une nouvelle vidéo en ligne, des dizaines de jeunes vidéastes proposent des liens renvoyant à leur chaîne dans un commentaire, invitant les internautes à venir la visiter. Il s'agit d'acquérir éventuellement davantage de notoriété. Ensuite, les chaînes de ces vidéastes moins populaires contiennent également des commentaires renvoyant à d'autres chaînes que nous avons consultées, et ainsi de suite.

Ce sont la mise en mots, la mise en actes et la mise en scène de l'intimité adolescente qui ont retenu notre attention dans le travail d'investigation, puis d'analyse du terrain ethnographique. Ces choix méthodologiques vont de pair avec la volonté d'avoir accès à des représentations et non à des expériences vécues de l'intimité. Dans ce contexte, le travail d'écriture, de réalisation et de stylisation de soi ne constitue pas une façon d'empêcher l'accès à une forme d'intimité « réelle », mais bien un élément d'analyse en soi. Si ce travail repose sur une forme de caricature de l'expérience adolescente, nous postulons que la caricature peut être un excellent outil pour comprendre les composantes de l'imaginaire social.

Après huit mois d'observation ethnographique sur YouTube, nous avons constitué un corpus de 80 vidéos, 40 réalisées par des garçons et 40, par des filles. Les vidéos ont été sélectionnées en lien avec les résultats issus du terrain d'immersion ethnographique. La référence à l'appartenance de genre étant apparue comme centrale dans les discours sur soi inscrits dans un format de confession intime sur YouTube, nous avons choisi des vidéos dans lesquelles les jeunes problématisent l'intimité en tant que garçon ou fille. L'objectif de recherche est alors axé sur les représentations du genre contenues dans ces vidéos que nous avons téléchargées puis transcrites.

Nous avons opté pour un mode de transcription intégrant les dimensions tant énonciatives que visuelles, les propos, mais aussi la tonalité et

le débit de la voix, la gestuelle, le décor, la coiffure, les vêtements, les indices du montage, etc. Les vidéos sont tournées le plus souvent dans l'intimité domestique, dans la chambre, sur le lit, avec en arrière-plan des éléments qui indiquent une appartenance de genre : petits lampions, bougies, coussins multicolores pour les filles ; guitare, ordinateur et jeu de fléchettes pour les garçons. La gestuelle est également très significative du point de vue de l'identification de genre. Tout un rituel est notamment observable dans la mise en scène de la féminité autour du toucher des cheveux. Tout en parlant, les jeunes filles les lissent avec les mains, les renvoient en arrière, les enroulent autour des doigts, etc. (Balleys, 2017b). Ces éléments ont donc occupé une place importante dans le travail de transcription afin de pouvoir être intégrés à l'analyse.

Nous avons appliqué ici une méthodologie empruntée au traitement d'entretiens collectifs, également réalisés à partir de contenus vidéo (Balleys, 2012 ; 2017a). L'extrait suivant, issu d'une vidéo intitulée « être en couple » et réalisée par un jeune Québécois de 14 ans, montre comment les éléments relatifs au travail de scénarisation prennent une place aussi importante que les *verbatim* dans la méthode de transcription :

Johan fait face à la caméra, en plan à la taille, debout dans la cuisine

« Bon là, je vais vous dire les avantages et les désavantages d'être en couple. Y a pas vraiment de désavantage d'être en couple, mais... ah ouais je le sais, parce qu'elle va te texter non-stop ! Attends-toi pas, si elle a un cell' ou un IPod, elle va te texter, mon ami ! Oh my God ! Pis les filles, juste pour dire, les filles qui sont en couple, là. Si votre chum ne vous répond pas, c'est pas mal sûr à 99.9 % qu'il est en train de jouer à [coupe, plan resserré sur son visage, il ouvre largement les bras et élève la voix] – *Call of Duty!* »

Nouvelle séquence. Mise en abyme.

De retour à la table de la cuisine, il personnifie cette fois une fille, porte un t-shirt sur la tête en guise de cheveux, a un téléphone intelligent entre les mains et écrit un texto :

« Allo *babe*. »

Il prend une voix aigüe et traînante, se touche les cheveux, le t-shirt.

« Je suis sûre qu'il va être content de mon message ! »

Changement de plan, il joue le petit ami, assis sur le canapé du salon, qui finit sa partie, manette en main :

« Oh je t'ai fait un 8 *shots*, man ! »

Il crie et effectue des mouvements de bassin mimant l'accouplement.

« *In your ass, bitch!* »

Retour vers la fille à sa table, avec son t-shirt en guise de chevelure.

« Pourquoi il me répond pas ? Hein ? Calisse... »

éclate en sanglots.

Retour vers le joueur, qui saute en l'air et s'énerve sur son jeu.

« Oh calisse, *fuck you, fuck you!* »

Retour sur le personnage de la fille, qui pleure bruyamment et se mouche dans son mouchoir. Elle en sort plusieurs de sous son t-shirt, comme si elle les prenait dans son soutien-gorge.

L'articulation entre matériau récolté et travail d'analyse

À partir de la multitude d'informations contenues dans les vidéos adolescentes sur YouTube, que faut-il observer, chercher et retenir ? En ce qui concerne notre enquête, nous avons concentré notre attention sur ce que Becker a nommé les « lignes de démarcation » visant à « séparer les choses en catégories » (2002, p. 239). Cette première volonté a été conjugée à une seconde : découvrir où se trouvent les éléments « d'identification à un ou plusieurs groupes » (Elias et Scotson, 1994, p. 184). Nous sommes parties à la recherche des discours permettant de « faire corps » avec une certaine entité, une catégorie d'individu ou un groupe clairement défini ou, au contraire, de « s'en distancier » et de marquer une différence dans une logique d'opposition. En résumé, nous avons dégagé les discours propres à créer des frontières entre un « nous » et un « eux » ou « elles ».

Par exemple, il est apparu assez rapidement au cours de l'observation ethnographique que les manifestations physiologiques du corps masculin pubère étaient fréquemment évoquées, voire mimées par les garçons : la masturbation, les érections matinales, l'apparition de la pilosité, la mue de la voix. Nous associons la problématisation de ces expériences intimes à la thématique du « grandir » et du « devenir un homme », ce qui est un élément explicite de fierté (Balleys, 2016b). En parlant de leurs pulsions sexuelles, les garçons se distinguent des « petits », mais aussi des filles qui, elles, sont systématiquement représentées comme des êtres dégoûtés, voire effrayés par la sexualité (*ibid.* ; Balleys, 2017b). La « ligne de démarcation » dont nous parle Becker (2002) est donc double dans ce contexte : elle dresse une frontière entre un « nous les gars » versus « elles les filles » et une autre

entre « nous les grands » versus « eux les petits ». Le travail de ritualisation de la masculinité adolescente opéré par les garçons pubères sur YouTube se concrétise à la fois par des énoncés et un répertoire symbolique comprenant une gestuelle, un décor, un débit de parole, une tenue vestimentaire, etc. Un principe d'opposition est perceptible dans les propos tenus par les jeunes, qui marquent des frontières claires entre ce qui s'apparente au vécu intime de « nous les gars » vis-à-vis de ce qui appartient au « nous les filles » : « on aime ça, nous, les filles » ; « on ne va pas se mentir, les mecs, hein ? » ; « si tu es une fille et que tu t'es reconnue » ; « il y a une question qu'on se pose nous, tous les mecs » ; « nous les hommes. Avec un grand H ». Face à la constitution du « nous » s'incarnent le « eux » et le « elles » qui représentent l'expérience de l'altérité. Les éléments scénographiques des vidéos complètent le contenu des discours. Les garçons, par exemple, se travestissent souvent dans les vidéos traitant de la relation de couple et dressent le portrait d'entités féminines leur étant en tout point distinctes : les filles sont dépeintes comme « contrôlantes », « jalouses » et « compliquées », alors qu'eux-mêmes se mettent en scène comme des êtres décomplexés et décontractés.

Les résultats ont ainsi montré que le genre constituait la principale délimitation sur laquelle la mise en scène de l'intimité adolescente était construite. L'appartenance de genre est mobilisée par les jeunes vidéastes afin de créer avec leur public un sentiment de proximité et de communauté. Pour obtenir une reconnaissance sociale sur YouTube, à savoir la « confirmation de leur valeur sociale » (Honneth, 2005, p. 45), des récits et des rituels sont construits en étant axés sur une intimité spécifiquement masculine ou féminine, selon un principe d'opposition : ce que les garçons adolescents expérimentent dans l'adolescence ne peut être appréhendé ni compris par les filles, et vice versa.

L'ajustement entre les questions de recherche et le type de données récoltées constitue la clé de voûte de l'enquête ethnographique sur Internet. YouTube est un formidable outil pour comprendre ce qui est aujourd'hui pourvoyeur de prestige et de valeur sociale au sein de la culture adolescente, puisque la pratique vidéo est axée sur une demande continue de reconnaissance. Pour saisir toute la richesse et la complexité des mises en scène de soi et de son intimité sur la plateforme, il est cependant nécessaire de situer la démarche dans le contexte de la socialisation adolescente, en particulier telle qu'elle se négocie au quotidien entre pairs.

Deux phases successives de méthodologie de recherche ont donc été nécessaires pour saisir le sens de ce format de mise en scène de l'intimité adolescente sur YouTube : l'observation non participante sur la plateforme réalisée à partir de critères de navigation précis et d'un cadre conceptuel portant sur la socialisation adolescente ; puis, la sélection et le traitement d'un corpus de vidéos constitué sur la base des premiers résultats issus de l'observation. Un véritable engagemement immersif dans la culture YouTube a été fondamental pour en comprendre les codes, le langage et les discours dominants. Plusieurs heures d'observation ont été effectuées par semaine, voire par jour selon les périodes. À partir des carnets ethnographiques tenus lors de ce premier volet de l'enquête, il a été possible de déterminer les critères de sélection puis de stabilisation d'un corpus de vidéos. Deux conditions ont été nécessaires à la production des résultats de cette recherche : premièrement, un cadre conceptuel délimité avec précision afin de circonscrire le terrain d'observation en ligne, et deuxièmement, une grande curiosité (voire une fascination) pour ces modes de présentation de soi médiatisés. Si l'approche sociologique vis-à-vis d'un objet n'est jamais neutre (Bourdieu, Chamboredon et Passeron, 1983), le visionnement de centaines de vidéos n'a jamais été vécu comme une corvée, mais bien comme un accès privilégié aux coulisses de la socialisation adolescente.

Alors que notre intérêt portait sur les positionnements identitaires des jeunes YouTubeurs, exprimés par la mise en scène d'une forme d'intimité, l'analyse du travail de ritualisation permettait de confirmer une distinction importante : l'enquête ne se préoccupe pas du vécu ou de l'expérience intime des garçons et filles, mais bien de la manière dont l'intimité est saisie sur YouTube comme ressource symbolique dans le but d'acquérir de la visibilité, du prestige et de la reconnaissance sociale. En effet, parler de son intimité sur YouTube, c'est apporter la preuve qu'on en possède une et que l'on est capable de la valoriser socialement. La question n'est donc pas de savoir si l'intimité ainsi performée est authentique ou réelle, mais de bien décrire les perceptions et les catégories sur lesquelles le travail de présentation de soi est construit, à partir d'un format de confession intime.

Notre posture méthodologique se situe à la croisée d'une relative fermeture dans la manière de délimiter l'objet observé et d'une volonté d'ouverture dans l'accueil des informations émanant du terrain. Alors

que l'appartenance de genre était peu présente dans notre hypothèse de départ, l'investigation ethnographique nous a permis de découvrir qu'elle était au cœur des revendications identitaires formulées par les jeunes sur YouTube. Les revendications d'appartenance de genre et la ritualisation du féminin et du masculin performées publiquement révèlent des représentations sociales qui, malgré leur caractère caricatural, voire hyperbolique, sont socialement signifiantes et productives. Lorsque des adolescents et des adolescentes décident de se mettre en scène sur YouTube, un modèle dominant se dégage clairement, construit sur un procédé de naturalisation du genre grâce à la distinction des expériences intimes féminines et masculines.

Références

- BALLEYS, Claire, « Entretiens en situation de réception collective et processus d'exclusion entre pairs adolescents », dans Philippe LE GUERN (dir.), *En quête de musique. Méthodologies de recherche à l'ère de la musimorphose*, Hermann, 2017a, p. 227-244.
- BALLEYS, Claire, « L'incontrôlable besoin de contrôle. Les performances de la féminité par les adolescentes sur YouTube », *Genre, Sexualité & Société*, vol. 17, 2017b.
- BALLEYS, Claire, « Gestion de l'intimité et affichage d'un territoire sentimental entre adolescents sur Internet », *Agora/Débats Jeunesses*, vol. 1, n° 72, 2016a, p. 7-19.
- BALLEYS, Claire, « "Nous les mecs." La mise en scène de l'intimité masculine sur YouTube », dans Éric DAGIRAL et Olivier MARTIN (dir.), *L'ordinaire d'Internet. Le Web dans nos pratiques et relations sociales*, Armand Colin, 2016b, p. 182-202.
- BALLEYS, Claire, *Grandir entre adolescents, à l'école et sur Internet*, Presses polytechniques et universitaires romandes, Lausanne, 2015.
- BALLEYS, Claire, *"Je t'aime plus que tout au monde". D'amitiés en amours, les processus de socialisation entre pairs adolescents*, thèse de doctorat, Faculté des sciences économiques et sociales, Université de Fribourg, 2012.
- BALLEYS, Claire et Sami COLL, « La mise en scène de la vie privée en ligne par les adolescents », *RESET – Recherches en sciences sociales sur Internet*, vol. 4, 2015, p. 1-18.
- BEAUD, Stéphane et Florence WEBER, *L'enquête de terrain. Produire et analyser des données ethnographiques*, La Découverte, 2003.
- BECKER, Howard S., *Les ficelles du métier*, La Découverte, 2002.
- BERGER, Peter et Thomas LUCKMANN, *La construction sociale de la réalité*, Armand Colin, 2008 [1966].
- BERRY, Vincent, « Ethnographie sur Internet: Rendre compte du "virtuel" », *Les Sciences de l'éducation – Pour l'ère nouvelle*, vol. 45, n° 4, 2012, p. 18-45.
- BEUSCART, Jean-Samuel et Kevin MELLET, « La conversion de la notoriété en ligne. Une étude des trajectoires de vidéastes pro-am », *Terrains et Travaux*, vol. 26, n° 1, 2015, p. 83-104.

- BOURDIEU, Pierre, Jean-Claude CHAMBOREDON et Jean-Claude PASSERON, *Le métier de sociologue*, Mouton, 1983.
- BURGESS, Jean et Joshua GREEN, *YouTube. Online Video and Participatory Culture*, Polity Press, 2018.
- CARON, Caroline, « Les jeunes et l'expérience participative en ligne », *Lien social et politiques*, n° 71, 2014, p. 13-30.
- CASTEL, Robert, « La face cachée de l'individu hypermoderne : l'individu par défaut », *Sociologie clinique*, vol. 1, 2006, p. 117-128.
- DAYAN, Daniel, « Les mystères de la réception », *Le Débat*, vol. 71, n° 4, 1992, p. 141-157.
- ELIAS, Norbert et John S. SCOTSON, *The Established and the Outsiders*, 2nd edition, SAGE Publications, 1994.
- GLEVAREC, Hervé, « L'attachement aux univers fictionnels et déplacement du champ du dicible et du pensable. La sériphilie des jeunes adultes », dans Martin JULIER-COSTES, Denis JEFFREY et Jocelyn LACHANCE (dir.), *Séries cultes et culte de la série chez les jeunes*, Presses de l'Université Laval, 2014.
- GOLD, Raymond I., « Jeux de rôle sur le terrain : observation et participation dans l'enquête sociologique », dans Daniel CEFALI (dir.), *L'enquête de terrain*, La Découverte, 2003.
- HINE, Christine, « Internet research as emergent practice », dans Patricia LEAVY et Sharlene Nagy HESSE-BIEBER (dir.), *Handbook of Emergent Methods*, Guilford Publications, 2008, p. 525-541.
- HONNETH, Axel, « Invisibilité : sur l'épistémologie de la « reconnaissance », *Réseaux*, vol. 129-130, n° 1, 2005, p. 39-57.
- JOUËT, Josiane et Coralie LE CAROFF, « L'observation ethnographique en ligne », dans Christine BARATS (dir.), *Manuel d'analyse du Web en sciences humaines et sociales*, Armand Colin, 2013, p. 147-165.
- LANGE, Patricia, *Kids on YouTube. Technical Identities and Digital Literacies*, Left Coast Press, 2014.
- LATZKO-TOTH, Guillaume et Madeleine PASTINELLI, « Par-delà la dichotomie public/privé : la mise en visibilité des pratiques numériques et ses enjeux éthiques », *tic& société*, vol. 7, n° 2, 2013.
- METTON-GAYON, Céline, *Les adolescents, leur téléphone et Internet : « tu viens sur MSN? »*, L'Harmattan, 2009.
- PASQUIER, Dominique, *Cultures lycéennes. La tyrannie de la majorité*, Autrement, 2005.
- PASQUIER, Dominique, *La culture des sentiments*, Éditions de la Maison des sciences de l'homme, 1999.
- RAUN, Tobias, *Out Online: Trans Self-Representation and Community Building on YouTube*, thèse de doctorat, Department of Culture and Identity, Roskilde University, 2012.
- ROTMAN, Dana et Jennifer PREECE, « The 'WeTube' in YouTube: Creating an online community through media sharing », *International Journal of Web-Based Communities*, vol. 6, n° 3, 2010, p. 313-333.

TROISIÈME PARTIE

Les méthodes sur traces

CHAPITRE 10

La densification des données : revaloriser la recherche qualitative à l'ère des données massives

*Guillaume Latzko-Toth, Claudine Bonneau,
Mélanie Millette*

Le « tournant computationnel » observé récemment en sciences sociales a provoqué une remise en question de la division du travail implicite entre les méthodologies qualitatives et quantitatives applicables à l'étude des médias (Berry, 2011). Avec l'accès à des ressources informatiques sans précédent ainsi qu'à de puissants algorithmes, les quantitativistes peuvent désormais surmonter certaines contraintes et s'attaquer à l'analyse des masses de données engendrées par l'usage des médias numériques.

L'expression *big data* (données massives) attire l'attention sur l'ampleur des ensembles de données, comme si la taille d'un échantillon était l'indicateur ultime de la fiabilité des conclusions qu'on peut tirer de son analyse alors que les études du genre sont souvent extrêmement hétérogènes (Kitchin et McArdle, 2016). Même si le volume ne constitue pas l'unique dimension des programmes de recherche axés sur les données massives, certains travaux manifestent une confiance empiriste (et positiviste) dans l'idée selon laquelle l'abondance de données offre un accès direct ou objectif à la « réalité sociale » (Kitchin, 2014 ; Manovich, 2012). Les intérêts que portent les scientifiques et les acteurs de l'industrie aux données massives convergent en effet dans l'objectif de parvenir à une modélisation à grande échelle des publics, des grappes d'utilisateurs et

d'utilisatrices, des tendances en matière de consommation et ainsi de suite¹.

Dans un contexte où les données massives gagnent en popularité, sans nécessairement que l'on en discute les biais ou les implications, il importe de réévaluer la contribution des méthodes qualitatives fondées sur de petits ensembles de données et leur complémentarité avec les approches quantitatives en ce qui concerne l'étude des dispositifs et des usages numériques. Une question guide ainsi notre réflexion : si l'on abandonnait la recherche portant sur de « petits » jeux de données, que perdrait-on ? Et à l'inverse, si l'on continue à mener des recherches sur de petits échantillons et corpus, comment devrait-on les analyser d'une manière qui soit à la fois significative et productive pour la recherche contemporaine en sciences humaines et sociales ?

La réalisation d'études sur les médias sociaux appelle des stratégies pour parvenir à recueillir un volume de données gérable tout en disposant de « suffisamment » d'information pour tirer des conclusions rigoureuses. Même lorsque le nombre de cas à l'étude est limité (par ex., 30 profils Facebook), la quantité de données numériques peut rapidement atteindre des proportions démesurées. Des mentions « J'aime » (Facebook) aux gazouillis (Twitter) en passant par les « épingles » (Pinterest) et les « snaps » (Snapchat), les interactions en ligne produisent un volume gigantesque de « traces numériques ».

Howison *et al.* définissent ces traces numériques comme « une archive de l'activité [...] constituée au moyen d'un système d'information en ligne [...], une preuve enregistrée confirmant qu'une chose s'est produite dans le passé » (2011, p. 769). Selon cette définition, l'enregistrement des activités en ligne apparaît comme un aspect essentiel de la trace numérique. Elle néglige toutefois les contenus générés par les usagers et usagères qui constituent également des traces (Bowker, 2007), mais d'un genre différent. Dans ce cas-ci, la production de contenu (écrire un gazouillis, publier une photo) est l'activité en soi, le contenu constituant une trace de cette

1. À titre d'exemple, la Computing Research Association, qui réunit des organisations provenant aussi bien du milieu universitaire que du monde de l'industrie, exerce des pressions pour que l'on augmente le budget de la National Science Foundation (NSF) de façon qu'elle puisse subventionner des projets de recherche en science des mégadonnées, sur la base des retombées attendues pour « le commerce, les sciences et la société » (Bryant *et al.*, 2008). Voir aussi Bollier (2010, p. 2022).

dernière avec les métadonnées qui l'accompagnent (par ex., auteur, date et heure, géolocalisation). Par conséquent, les traces numériques comprennent toutes les données associées à une personne et appréhendables sur les diverses plateformes et applications en ligne. Comme le soulignent Howinson *et al.* (2011), la promesse des traces numériques réside dans la possibilité d'un accès « direct » aux pratiques « réelles », grâce à la capture et à l'analyse des traces qu'elles laissent en ligne. Cela répondrait à une limite des méthodes ethnographiques traditionnelles qui offrent des interprétations à partir d'un nombre limité d'observations ou de comptes rendus déclaratifs sur les pratiques, lesquels font appel à la mémoire des personnes participantes (et à leur honnêteté). Néanmoins, comme le font remarquer certaines critiques, l'interprétation de ces masses de données peut s'avérer très difficile (Kitchin, 2014), en particulier si l'on tient compte de l'importance du contexte dans l'étude des pratiques médiatiques (boyd et Crawford, 2012 ; Quan-Haase *et al.*, 2015). Dans les recherches sur les médias sociaux fondées sur des ensembles massifs de traces, la quantité tend à être considérée comme synonyme de qualité, ce qui détourne l'attention de questions sérieuses sur la représentativité et la validité de l'échantillon (Brooker *et al.*, 2016 ; Tufekci, 2014). Par ailleurs, la prolifération des traces numériques pose un défi de taille à leur analyse qualitative au moyen des méthodes manuelles, sans parler du temps requis afin de mener une telle entreprise.

Une solution consiste à réduire l'« étendue » des données (le nombre de points de données ou de sujets étudiés, par exemple) tout en améliorant leur « profondeur », c'est-à-dire la « densité » de chaque point. C'est ce que nous appelons « densifier » les données. Avant de proposer une définition plus rigoureuse de la notion de « donnée dense », nous employons ici une métaphore inspirée de la géologie pour illustrer ce que signifie troquer l'étendue pour la densité et pour montrer en quoi cette stratégie peut être féconde. Si l'on considère chaque point ou cas étudié comme un échantillon de forage prélevé dans le sol, le fait de forer plus profondément permet alors une analyse plus complète des interactions complexes entre les différentes couches du terrain. En supposant que la multiplication des forages n'améliore pas de manière significative, au-delà d'un certain nombre, notre compréhension des choses – ce que l'on appelle la « saturation » en recherche qualitative (Bowen, 2008) –, on peut alors envisager de compenser la perte de représentativité territoriale que permettrait un

grand nombre de prélèvements de sol par les connaissances approfondies obtenues grâce à un forage plus profond, mais sur un nombre restreint de points dans le site. L'idée est donc de miser sur une compréhension en profondeur du terrain plutôt que sur une représentativité en surface.

Nous avançons qu'il est possible de bonifier les méthodes qualitatives utilisées pour étudier les médias sociaux et leurs usages en densifiant les données recueillies sur les plateformes. Nous expliquons notre approche de densification en trois couches qui correspondent respectivement à la contextualisation des données, à leur description fine et à la signification des données issues des traces en ligne. Puis, nous illustrons cette approche en présentant trois stratégies de densification mobilisées dans le cadre d'études qualitatives sur les usages des médias sociaux : les entrevues sur traces, l'observation transplateforme et la collecte manuelle². Ces stratégies présentent plusieurs avantages pour la recherche qualitative tout en faisant ressortir leur complémentarité avec d'autres méthodes, notamment celles associées aux données massives.

Les « données denses »

Le terme « données denses » est apparu dans la littérature autour des années 1990, vraisemblablement sous l'influence exercée par les travaux de l'anthropologue Clifford Geertz dans les domaines de l'ethnographie et de la méthodologie qualitative. On associe les données denses au travail ethnographique donnant lieu à des descriptions détaillées et étoffées des pratiques culturelles (Geertz, 1973). Dans une introduction sur la méthode des études de cas, Stake fait référence à Geertz en soulignant que « les données issues des études de cas ont été qualifiées de “denses”, ce qui ne signifie pas forcément qu'elles forment des piles hautes, mais plutôt qu'elles ont “une texture semblable à celle de spaghettis densément enchevêtrés” » (traduction libre, 1985, p. 279). Cette expression en évoque une autre, celle de « données riches », employée par Howard Becker (1970) pour désigner des données sociologiques denses, hautement texturées et contextualisées. Dans une réflexion sur la validité des recherches qualitatives, Onwuegbuzie et Leech soulignent que « l'un des moyens essentiels de conférer de la

2. Ces stratégies sont présentées de manière plus détaillée et illustrées par des études de cas dans les chapitres suivants de Gallant *et al.* (entrevue sur traces) et Bonneau (collecte manuelle) au sein de cet ouvrage.

crédibilité à des conclusions consiste à recueillir des données riches et denses, c'est-à-dire des données suffisamment détaillées et complètes pour maximiser la capacité d'en tirer un sens » (traduction libre, 2007, p. 244).

Plus récemment, Wang (2013 ; 2016) emploie le terme « données denses » et fait un clin d'œil à Clifford Geertz dans sa critique de l'enthousiasme que suscite la recherche sur les données massives. À ses yeux, l'expression incarne la pertinence durable des méthodes ethnographiques qualitatives « en cette ère de données massives » (traduction libre de Wang, 2016). La chercheuse avance que « les données denses sont tout le contraire des données massives » (*ibid.*), car elles se forment à partir d'une substance difficile à quantifier comme des émotions, des récits, des manières de voir le monde dont on aplanirait les significations en « les normalisant, en les uniformisant, en les définissant [et] en les regroupant » (*ibid.*) si l'on voulait les organiser dans des jeux massifs de données afin d'en tirer des généralisations. Les données denses proviennent d'échantillons de petite taille, explique Wang, mais elles offrent néanmoins « une profondeur d'analyse incroyable en ce qui a trait au sens et aux récits » (*ibid.*).

Notre approche des données denses provient également de l'ethnographie virtuelle et de la netnographie (Hine, 2000 ; 2015 ; Kozinets, 2009). Ces méthodes produisent des ensembles de données qualitatives prenant différentes formes, comme des notes approfondies de terrain, des transcriptions d'entrevue et une connaissance intime de la culture étudiée, de sa langue, de ses rites et de ses symboles.

La distinction entre données denses et données massives

Il nous apparaît que les données denses sont à la fois complémentaires et une solution de remplacement aux approches axées sur les données massives, et ce, sur plusieurs plans. La première différence concerne l'échelle, soit l'usage de « petits ensembles de données ». Par cette expression, nous entendons un nombre relativement réduit d'éléments ou de cas (un petit N) à partir desquels collecter des données, de telle sorte qu'il soit possible pour une personne ou pour une petite équipe de recherche d'en faire la codification manuelle en recourant peu aux algorithmes et aux outils de modélisation, contrairement aux stratégies qui nécessitent un recours systématique aux processus automatisés de collecte et d'analyse de données massives. Cela dit, comme le rappelle la dimension « massive » des

données, la notion de « petit N » reste floue dans l'absolu et ne peut se préciser qu'en fonction du contexte et de la nature des données. À titre d'exemple, on considère généralement qu'un corpus d'entretiens semi-dirigés auprès d'un groupe dépassant la centaine de sujets constitue un vaste échantillon, alors qu'un corpus de 100 messages dans Twitter sera considéré comme un petit échantillon. Par conséquent, la taille de l'échantillon doit être évaluée par rapport à la capacité à le gérer, ce qui fait implicitement référence à la quantité de travail humain et de temps nécessaire pour réaliser l'analyse. Elle est influencée en outre par les normes établies en matière de représentativité statistique par les épistémologies positivistes. Si un réseau social comprend des centaines de millions de membres, un échantillon de quelques centaines de ces personnes sera sans doute jugé trop restreint pour être pertinent du point de vue statistique³. Les petits ensembles de données sont ainsi trop limités pour être représentatifs au sens statistique du terme et « suffisamment petits » pour pouvoir être traités par une équipe restreinte d'analystes dans le but de produire une représentation exhaustive d'un phénomène situé et complexe.

Une autre différence fondamentale entre données denses et données massives a trait aux questions de recherche auxquelles leur analyse permet de répondre. Alors que les données massives sont utiles pour l'étude des liens entre les individus, l'établissement de grappes thématiques et les tendances à grande échelle, les données denses visent à saisir la spécificité des usages, les subtilités des textes et de leurs connotations ainsi que le sens que leur donnent les individus. Comme l'explique Marwick,

[il] peut être utile de cerner de grandes tendances, mais on risque aussi de négliger la façon dont les usages s'implantent [dans les médias sociaux], les raisons pour lesquelles ils s'établissent ainsi que l'interprétation qu'en font les usagers. Les études quantitatives servent souvent à déterminer les liens et les réseaux et à les interpréter « objectivement » *a posteriori* sur la base de statistiques et de chiffres. La recherche qualitative vise plutôt à analyser la construction même du sens, en situant l'usage de la technologie dans un contexte social, un lieu et une époque en particulier (traduction libre de Marwick, 2013, p. 119).

3. De toute façon, le but ultime des recherches fondées sur de petits échantillons n'est pas de produire des résultats généralisables, mais plutôt d'analyser des enjeux humains de nature complexe.

Le modèle de densification des données

Les données sont toujours le résultat d'une construction, qu'elles soient massives ou petites, et chaque type de données favorise certaines formes d'analyses et en limite d'autres (Gitelman et Jackson, 2013 ; Luka et Millette, 2018). La « densification » des données se rapporte à un aspect de ce processus de construction qui consiste à compléter les données par des informations richement texturées ou, en d'autres termes, « en leur ajoutant des couches ». À l'image d'un oignon, il s'agit « d'enrober » les données et de les enrichir de plusieurs couches de métadonnées – au sens littéral d'empiler donnée sur donnée.

Le modèle de densification des données s'inscrit dans le paradigme interprétatif/constructiviste de la recherche qualitative, suivant lequel un phénomène social doit être appréhendé (1) dans son contexte ; (2) au moyen de descriptions fines et (3) à la lumière du sens que les personnes donnent à leurs propres actions (Geertz, 1973 ; Lincoln et Guba, 1985 ; Schütz, 1967). En reprenant cette idée, nous proposons donc ici un modèle de densification des données comportant trois couches⁴.

1. Contextualisation : Dans une recherche qualitative sur les médias sociaux à partir d'un petit ensemble de données, la « première couche » est constituée d'informations contextuelles. Il s'agit de comprendre les circonstances dans lesquelles la pratique en ligne survient, ainsi que les affordances techniques et les conventions culturelles qui les façonnent (Hine, 2015).
2. Description : On ajoute ensuite une « deuxième couche » en enrichissant les traces recueillies de « descriptions denses » des pratiques qui les ont générées. On peut produire ce genre de descriptions par un engagement actif auprès des personnes étudiées et de l'environnement numérique dont il s'agit, ainsi qu'avec les données elles-mêmes, par exemple à l'aide d'une observation en ligne.
3. Signification : Finalement, et en cohérence avec les objectifs d'une recherche qui porte sur un phénomène social, les pratiques numériques

4. Nous sommes conscients qu'il est possible de complexifier et de raffiner le modèle proposé. On pourrait envisager la possibilité d'autres couches ou les ordonner différemment. Notre but n'est pas d'offrir un modèle définitif et exhaustif, mais d'insister sur l'utilité d'ajouter aux données des « couches de densité ».

ne peuvent être étudiées sans tenir compte des perspectives multiples des acteurs et des processus de création du sens (boyd et Crawford, 2012). Par conséquent, il faut prévoir une « troisième couche » qui saisisse les expériences des usagers et usagères, et le sens accordé à leurs pratiques. Le contenu de cette couche découlera des interactions directes avec les sujets dont on étudie les traces afin d'expliquer la compréhension de leurs interventions en ligne et de la valeur qu'ils y accordent⁵.

Chacune de ces couches (contextualisation, description, signification) peut être divisée en sous-couches destinées à répondre à des objectifs de recherche précis. Elles peuvent également donner forme à des types de données et recourir à des méthodes d'analyse variées. Par exemple, une équipe qui s'intéresse à la présentation de soi sur Instagram pourrait enrichir la deuxième couche (description) en ajoutant des détails sur l'organisation visuelle des profils en ligne ou encore en offrant une analyse sémiotique des photos publiées par les usagers.

À ce titre, nous ne considérons pas les données comme des entités neutres, encore moins prêtes à l'emploi. En tant que chercheurs et chercheuses ayant un point de vue situé, nous devons prendre conscience des cadrages que nous mobilisons pour construire nos données (Gitelman et Jackson, 2013; Markham, 2013; Luka et Millette, 2018). En effet, les données ne sont jamais une médiation directe d'une réalité étudiée; il faut plutôt « les cuisiner avec soin », c'est-à-dire les affiner et les contextualiser avant qu'elles puissent être utiles et nous renseigner (traduction libre de Bowker, 2005, p. 183-184).

Trois applications du modèle de densification des données

Plusieurs méthodes ou combinaisons de méthodes peuvent permettre de densifier les données. Dans cette section, nous proposons trois stratégies méthodologiques pour appliquer le modèle de densification des données dans le cadre de recherches qualitatives portant sur des traces d'usage dans les médias sociaux. Ces stratégies sont adaptables et peuvent être

5. Cela fait écho aux techniques de « vérification par les membres », par lesquelles les chercheurs confirment la validité de leur interprétation des données auprès des participants (Lincoln et Guba, 1985).

combinées entre elles ou avec d'autres méthodes afin d'étudier différentes pratiques et traces numériques. Il ne s'agit pas de stipuler que ces méthodes sont les seules menant à la densification des données, mais plutôt d'exemplifier l'application.

L'entrevue sur traces numériques et la visite commentée

Le premier exemple est une recherche dans laquelle la méthode de l'«entrevue sur traces» a été choisie. Il s'agit d'un entretien pendant lequel on invite les usagers à réfléchir sur leurs propres traces numériques et à formuler ainsi un métadiscours les concernant. Dans une certaine mesure, on peut parler d'une forme de coanalyse réalisée avec le sujet dont on examine les traces (commentaires, mentions «j'aime», etc.). Ce genre d'entretien peut prendre différentes formes. Dubois et Ford (2015), par exemple, recourent à un protocole en deux temps : les traces sont captées et traitées quantitativement pour générer des visualisations qui sont ensuite présentées aux personnes afin d'obtenir leurs commentaires. Avec cette méthode, toutefois, les traces ne font pas, pour l'essentiel, directement l'objet d'une discussion avec les personnes participantes. De plus, les visualisations orientent l'attention sur certains aspects tout en négligeant d'autres éléments. Malgré tout, elles sont utiles pour lancer l'entretien (*ibid.*). Une variante consiste à se servir d'un «tableau de bord» présentant des statistiques tirées du compte de l'utilisateur ou de l'utilisatrice (par ex., Twitonomy).

Une approche différente, la visite commentée des traces, consiste à examiner une sélection de traces réelles avec la personne les ayant émises (Gallant *et al.*, le présent ouvrage). Reposant sur une logique d'exploration, cette approche exige une sélection minutieusement réfléchie des traces à commenter afin d'éviter que l'entretien ne soit trop long. Préférentiellement, les entretiens sont enregistrés sur vidéo au moyen d'un logiciel de capture d'écran couplé à un enregistrement audio (certains logiciels comme Camtasia permettent la capture simultanée de l'activité à l'écran et du son ambiant). Bref, la «visite commentée» offre une méthode de collecte et de densification des données comportant une exploration et une capture sélectives des traces numériques dans leur contexte d'origine en présence de la personne qui les a produites. Le dialogue concomitant sur les traces, mené entre la personne effectuant l'enquête et la personne participante,

fournit la matière qui sert ensuite à densifier les données en y ajoutant les couches relatives au contexte, à la description et au sens.

La collecte manuelle de données

Dans le cadre d'une autre recherche, la « collecte manuelle de données » a été sélectionnée pour densifier des données. Cette méthode consiste à recueillir manuellement des traces numériques, en particulier les messages (par ex., les gazouillis sur Twitter), les images (par ex., les publications sur Instagram) et d'autres types de contenus produits à travers l'usage. Cette approche est particulièrement fructueuse dans les cas où les limites du « terrain » sont floues, fluides et situées (Bonneau, le présent ouvrage). Cette situation survient dans les cas où l'étude n'est pas circonscrite à une communauté bien définie ou encore lorsqu'il n'est pas possible de prédéfinir des critères de recherche des données sur la base d'un profil sociodémographique, d'un sujet ou d'un événement particulier. Dans pareil cas, il est impossible de recourir à des fonctions d'exploration du texte en cherchant à atteindre une cible mouvante qui ne peut être délimitée à l'avance. Ces difficultés demandent une approche souple permettant de constituer un corpus grâce au bricolage et à la sérendipité. La collecte manuelle présente alors un avantage supplémentaire, celui d'encourager l'exploration et de permettre de se familiariser avec le format « natif » des traces, c'est-à-dire la configuration d'origine dans laquelle les traces se présentent.

Dans une étude sur Twitter en lien avec le travail, Bonneau et Sergi (2015) ont tenu un journal des messages recueillis, assortis de métadonnées comme les URL. Les chercheuses se sont appuyées sur ces métadonnées pour élaborer des comptes rendus détaillés sur chacune des traces recueillies, densifiant ainsi la couche « description ». En examinant le profil des usagers et usagères et leurs messages antérieurs, elles sont parvenues à inclure de nouveaux contextes dans lesquels se déroulent différents types de pratiques de « travail à haute voix », ce qui a permis d'ajouter des éléments à la couche de contextualisation. La couche des significations a été reportée à une autre étape de la recherche, qui comportera des entretiens.

L'observation transplateforme de longue durée

Enfin, le troisième exemple d'application du modèle de densification relève d'une méthodologie empruntant aux principes ethnographiques⁶, l'« observation transplateforme de longue durée ». Cette stratégie repose sur l'observation prolongée des phénomènes en ligne conjugée à l'agilité de suivre les usagers d'une plateforme à l'autre (de Twitter aux blogues, des blogues à Facebook, etc.). Traditionnellement, les ethnographes séjournent sur un terrain pendant une période prolongée et recueillent une vaste quantité de données sur la langue, la culture, les symboles et les valeurs d'un groupe social. De manière similaire, la méthode d'observation proposée ici implique une immersion de plusieurs mois ou même de quelques années et contribue grandement à approfondir ce type de connaissances.

Les internautes communiquent bien souvent en adoptant des pratiques transplateformes (Millette, 2013), par exemple en publiant un message dans Twitter à propos d'un billet de blogue ou en publiant un commentaire sur Facebook à propos d'une nouvelle lue ailleurs. L'agilité implique une certaine souplesse dans la manière de rassembler les données et la capacité de suivre les internautes lorsqu'ils passent d'une plateforme à une autre. Ainsi, le protocole méthodologique d'une recherche qui examine les usages de Twitter pourrait comprendre l'observation d'autres sites. Cette agilité transplateforme a permis la production d'une analyse détaillée de la baladodiffusion indépendante à Montréal (Millette, 2011) et des usages des médias sociaux par les minorités franco-canadiennes (Millette, 2015). Dans ces deux recherches, cette stratégie d'observation transplateforme de longue durée a grandement servi à densifier les couches de contextualisation et de description des données (les pratiques et les traces), tandis que des entretiens ont permis à la chercheuse de dégager le sens accordé aux pratiques par les acteurs.

* * *

6. Voir, par exemple, les techniques « d'engagement prolongé » et « d'observation persistante » proposées par Lincoln et Guba (1985) dans le but d'assurer la fiabilité en recherche qualitative. Sur la souplesse nécessaire à la définition des limites du domaine, consulter Green (1999).

Rejetant la simple affirmation selon laquelle la recherche menée sur de petits ensembles de données est dépassée, ce chapitre propose des stratégies concrètes pouvant servir à « densifier » les données provenant de petits corpus et à favoriser une analyse en profondeur des phénomènes à l'étude, et ce, sans remettre en question les perspectives inédites qu'ouvrent les processus automatisés de collecte et d'analyse des données massives.

Nous estimons que le modèle de densification des données est porteur pour la recherche qualitative sur le numérique. Fondé sur l'enrichissement de trois « couches » constituées de données secondaires, les couches de description, de contextualisation et de signification, ce modèle ne repose pas sur une démarche unidirectionnelle, mais bien relationnelle. La densification se produit lorsque des liens s'établissent entre plusieurs sources de données, permettant une appréhension complexe d'un phénomène. La densité croît non seulement du côté des traces numériques, mais aussi des autres données qualitatives qui sont recueillies en chemin. À titre d'exemple, il est tout aussi possible de conférer une « épaisseur » aux données d'entretien par des références aux traces des sujets, que de densifier des traces en s'appuyant sur les observations émanant des entretiens.

Notre intention n'est pas d'avancer qu'il faut privilégier les échantillons de petite taille dans la recherche sur les médias sociaux. Les projets fondés sur une petite quantité de données ne peuvent pas conduire aux résultats robustes et généralisables auxquels peuvent aboutir certaines études quantitatives à vaste échelle. Il ne faut pas non plus voir dans les stratégies de densification un remède aux limites inhérentes aux méthodologies axées sur les petits ensembles de données. Les données denses sont plutôt à considérer en complémentarité avec les données massives, ce qui constitue en soi un argument en faveur de la collaboration entre les deux approches et le développement de devis de recherches mixtes (Ford, 2014 ; Lewis *et al.*, 2013). Quant au modèle de densification proposé, il permet de systématiser le travail méthodologique sur de petits jeux de données et de cerner un maximum de dimensions relatives à un phénomène.

Références

- BECKER, Howard S., *Sociological Work: Method and Substance*, Transaction Books, 1970.
- BERRY, David M., « The computational turn: Thinking about the digital humanities », *Culture Machine*, vol. 12, 2011, p. 1-22.

- BOLLIER, David, *The Promise and Peril of Big Data*, The Aspen Institute, 2010.
- BONNEAU, Claudine et Viviane SERGI, « Travailler à haute voix sur Twitter : Raconter son travail pour l'organiser », dans *Actes du colloque international Org&Co*, groupe Org&Co et CERTOP, 2015, p. 131-141.
- BOWEN, Glenn A., « Naturalistic inquiry and the saturation concept: A research note », *Qualitative Research*, vol. 8, n° 1, 2008, p. 137-152.
- BOWKER, Geoffrey C., « The Past and the Internet », dans Joe KARAGANIS (dir.), *Structures of Participation in Digital Culture*, Social Science Research Council, 2007, p. 20-36.
- BOWKER, Geoffrey C., *Memory Practices in the Sciences*, The MIT Press, 2005.
- BOYD, danah et Kate CRAWFORD, « Critical questions for big data: Provocations for a cultural, technological and scholarly phenomenon », *Information, Communication & Society*, vol. 15, n° 5, 2012, p. 662-679.
- BROOKER, Phillip, Julie BARNETT, Timothy CRIBBIN et Sanjay SHARMA, « Have we even solved the first "Big Data challenge?" Practical issues concerning data collection and visual representation for social media analytics », dans Helene SNEE, Christine HINE, Yvette MOREY, Steven ROBERTS et Hayley WATSON (dir.), *Digital Methods for Social Science. An Interdisciplinary Guide to Research Innovation*, Palgrave Macmillan, 2016, p. 34-50.
- BRYANT, Randal E., Randy H. KATZ et Edward D. LAZOWSKA, *Big-Data Computing: Creating Revolutionary Breakthroughs in Commerce, Science, and Society*, rapport préparé pour le Computing Community Consortium committee de la Computing Research Association, 2008.
- DUBOIS, Elizabeth et Heather FORD, « Trace interviews: An actor-centered approach », *International Journal of Communication*, vol. 9, n° 25, 2015, p. 2067-2091.
- FORD, Heather, « Big data and small: Collaborations between ethnographers and data scientists », *Big Data & Society*, vol. 1, n° 2, 2014, p. 1-3.
- GEERTZ, Clifford, *The Interpretation of Cultures: Selected Essays*, Basic Books, 1973.
- GITELMAN, Lisa et Virginia JACKSON, « Introduction », dans Lisa GITELMAN (dir.), *"Raw Data" is an Oxymoron*, The MIT Press, 2013, p. 1-14.
- GREEN, Nicola, « Disrupting the field: Virtual reality technologies and "multisited" ethnographic methods », *American Behavioral Scientist*, vol. 43, n° 3, 1999, p. 409-421.
- HINE, Christine, *Ethnography for the Internet: Embedded, Embodied and Everyday*, Bloomsbury Publishing, 2015.
- HINE, Christine, *Virtual Ethnography*, SAGE Publications, 2000.
- HOWISON, James, Andrea WIGGINS et Kevin CROWSTON, « Validity issues in the use of social network analysis with digital trace data », *Journal of the Association for Information Systems*, vol. 12, n° 12, 2011, p. 767-797.
- KITCHIN, Rob, « Big Data, new epistemologies and paradigm shifts », *Big Data & Society*, vol. 1, n° 1, 2014, p. 1-12.
- KITCHIN, Rob et Gavin MCARDLE, « What makes Big Data, Big Data? Exploring the ontological characteristics of 26 datasets », *Big Data & Society*, vol. 3, n° 1, 2016, p. 1-10.

- KOZINETS, Robert V., *Netnography: Doing Ethnographic Research Online*, SAGE Publications, 2009.
- LEWIS, Seth C., Rodrigo ZAMITH et Alfred HERMIDA, « Content analysis in an era of Big Data: a hybrid approach to computational and manual methods », *Journal of Broadcasting & Electronic Media*, vol. 57, n° 1, 2013, p. 34-52.
- LINCOLN, Yvonna S. et Egon G. GUBA, *Naturalistic Inquiry*, SAGE Publications, 1985.
- LUKA, Mary Elizabeth et Mélanie MILLETTE, « (Re)framing Big Data: Activating situated knowledges and a feminist ethics of care in social media research », *Social Media + Society*, vol. 4, n° 2, 2018.
- MANOVICH, Lev, « Trending: The promises and challenges of big social data », dans Matthew K. GOLD (dir.), *Debates in the Digital Humanities*, University of Minnesota Press, 2012, p. 460-475.
- MARKHAM, Annette N., « Undermining “data”: A critical examination of a core term in scientific inquiry », *First Monday*, vol. 18, n° 10, 2013.
- MARWICK, Alice E., « Ethnographic and qualitative research on Twitter », dans Katrin WELLER, Axel BRUNS, Jean BURGESS, Merja MAHRT et Cornelius PUSCHMANN (dir.), *Twitter and Society*, Peter Lang, 2013, p. 109-122.
- MILLETTE, Mélanie, *L'usage des médias sociaux dans les luttes pour la visibilité: le cas des minorités francophones au Canada anglais*, thèse de doctorat en communication, Université du Québec à Montréal, Montréal, 2015.
- MILLETTE, Mélanie, « Pratiques transplateformes et convergence dans les usages des médias sociaux », *Communication et organisation*, n° 43, 2013, p. 47-58.
- MILLETTE, Mélanie, « Independent podcasting as a specific online participative sub-culture: A case study of Montreal's podcasters », *Selected Papers of Internet Research – AOIR*, vol. 1, 2011.
- ONWUEGBUZIE, Anthony J. et Nancy L. LEECH, « Validity and qualitative research: An oxymoron? », *Quality & Quantity*, vol. 41, n° 2, 2007, p. 233-249.
- QUAN-HAASE, Anabel, Kim MARTIN et Lori McCAY-PEET, « Networks of digital humanities scholars: The informational and social uses and gratifications of Twitter », *Big Data & Society*, vol. 2, n° 1, 2015, p. 1-12.
- SCHÜTZ, Alfred, *The Phenomenology of the Social World*, Northwestern University Press, 1967.
- STAKE, Robert E., « Case study », dans John NISBET, Jacquetta MEGARRY et Stanley NISBET (dir.), *World Yearbook of Education 1985: Research, Policy and Practice*, Kogan Page, 1985, p. 277-285.
- TUFERCI, Zeynep, « Big questions for social media Big Data: Representativeness, validity and other methodological pitfalls », dans *ICWSM 14: Proceedings of the 8th International AAAI Conference on Weblogs and Social Media*, 2014.
- WANG, Tricia, « Why Big Data needs thick data », *Ethnography Matters*, 2016.
- WANG, Tricia, « Big Data needs thick data », *Ethnography Matters*, 2013.

CHAPITRE 11

La visite commentée : documenter les pratiques numériques par l'entretien sur traces

*Nicole Gallant, Katherine Labrecque,
Guillaume Latzko-Toth, Madeleine Pastinelli*

La « visite commentée » est une technique de collecte de données qui se situe dans le courant méthodologique de l'analyse des « traces » numériques – l'étude des pratiques numériques par l'intermédiaire des traces laissées, volontairement ou non, par les individus sur Internet. Elle est illustrée ici à partir de deux projets de recherche¹.

Une portion substantielle de la recherche universitaire sur les traces numériques consiste à « aspirer » un important volume de contenu disponible sur les plateformes des principaux médias socionumériques, en particulier des *tweets* (boyd, 2010; Hong, Convertino et Chi, 2011; Millette, Millette et Proulx, 2012). Bien que certaines recherches s'appuient sur un méticuleux mais fastidieux codage à la main (Millette, 2014), l'examen de ces traces fait le plus souvent l'objet d'un codage automatisé, allant du simple décompte du nombre de traces ou d'occurrences de mots jusqu'à des visualisations sous forme d'agrégats, en passant par des fouilles de texte (*text mining*) ou de codage automatique des informations de géolocalisation (par

1. La mise au point de cette méthode n'aurait pas été possible sans l'aide, pour le projet initial, de nos quatre assistants de recherche et enquêteurs, Olivier Gadeau, Pierre-Élie Hupé, Frédéric Nadeau et Guillaume Tremblay-Boily, que nous remercions pour leur travail rigoureux et leurs commentaires précieux.

exemple, Bruns et Burgess, 2012). Quel que soit le mode de codage adopté, toutefois, ces techniques comportent certaines limites en termes de contextualisation des traces et de compréhension des significations subjectives des pratiques numériques.

En réponse à ces limites, plusieurs autres pratiques de collecte plus participatives ont été mises au point, en suscitant une interaction tripartite entre les équipes de recherche, les traces elles-mêmes et les personnes à l'origine de ces traces, grâce à des entretiens dits « sur traces » (Dubois et Ford, 2015). C'est dans cette perspective que nous avons conçu en 2012 la technique de collecte de données présentée ici – la « visite commentée » –, qui engage les personnes participantes dans l'observation et l'analyse de leurs traces dans une visée interprétative. Notre technique de « visite commentée » permet non seulement de récolter des traces, mais aussi de les observer en contexte, tout en bénéficiant de la réflexion des individus sur leurs propres activités en ligne.

Depuis, et en parallèle, plusieurs techniques similaires ont été élaborées. Ces techniques peuvent se répartir en deux grands champs. Le premier vise l'analyse des traces elles-mêmes, que le commentaire de la personne enrichit. C'est le cas notamment du *media go-along* qui combine une entrevue classique avec une visite verbale et visuelle des plateformes numériques, comme le fait Moller Jorgensen (2016) avec des applications de rencontres. Le second champ d'utilisation de ces techniques consiste plutôt à utiliser la visite des traces comme une sorte de stimulus visant à faire émerger un discours sur les pratiques numériques. Dans ces cas, c'est le discours qui sert de principal support à l'analyse, laquelle porte sur les représentations des applications et des plateformes Internet, plutôt que sur les traces que les individus y laissent (McVeigh-Schultz et Baym, 2015). Par exemple, la *scrollback method* consiste à revenir en arrière sur Facebook et à « coanalyser » avec les individus l'évolution de la construction de leur identité en ligne à partir de leurs propres traces numériques, mais sans nécessairement enregistrer celles-ci (Robards et Lincoln, 2017). Ce type de technique peut être utilisé pour des situations qui ne laissent en fait aucune trace visible de leur navigation, telle la méthode d'auto-confrontation (Beauvisage, 2016 ; voir Pélissier, le présent ouvrage).

Pour sa part, la technique de « la visite commentée » que nous avons mise au point permet les deux types d'utilisation, c'est-à-dire à la fois l'analyse des traces ou l'utilisation des traces comme stimuli.

La technique de la visite commentée

Bien que la visite commentée puisse être employée de manières diverses selon les objectifs de recherche, le cœur de la technique repose sur un déroulement type, composé d'une rencontre individuelle en deux parties. La première prend la forme d'un entretien semi-dirigé classique et la seconde se déroule devant un ordinateur, alors que l'individu commente les traces de sa propre activité numérique tandis qu'elles sont observées à l'écran et capturées par un enregistrement vidéo dynamique de l'écran, grâce à un logiciel tel que Camtasia².

La visite commentée permet une exploration des traces d'activité publiques ou privées sur un ou plusieurs sites Internet³ en les enrichissant d'informations et d'explications verbales liées au contexte de la production de la trace, contrairement aux enquêtes qui «aspirent» ces traces avec un dispositif informatique permettant leur collecte exhaustive, mais en les détachant des représentations qu'en a l'individu et des significations qu'il leur prête. La visite commentée permet aussi la saisie de la trace sous une forme relativement fidèle à sa présentation au moment où l'individu l'a produite (voir figure 11.1).

FIGURE 11.1

Exemple de trace saisie dans son contexte de présentation (capture d'une vidéo Camtasia)



2. L'enregistrement du son ambiant (au moyen du micro de l'ordinateur) permet de disposer d'un enregistrement sonore parfaitement synchrone avec l'activité qui se déroule à l'écran.

3. On peut par exemple se connecter au compte d'un individu sur une plateforme telle que Facebook et explorer les traces (partages, commentaires, discussions, mentions «j'aime», etc.) archivées dans son historique personnel.

La visite commentée : deux exemples de recherche

Pour illustrer les stratégies concrètes mises en œuvre pour la collecte et l'analyse de traces dans le cadre de la visite commentée, nous nous pencherons sur deux projets. Le premier illustre la fécondité de la technique pour l'étude des pratiques numériques elles-mêmes et sur une plateforme numérique précise, alors que le second montre l'utilité de la méthode pour explorer une diversité de sites dans le cadre d'une étude sur des pratiques sociales plus larges.

Le projet Printemps érable

Notre méthode de « visite commentée » a initialement été développée expressément pour un projet portant sur la circulation et la réception de l'information sur Facebook pendant la grève étudiante de 2012 au Québec (Gallant *et al.*, 2015)⁴. Nous souhaitons observer de façon détaillée ce qui s'était effectivement passé en ligne, tout en tenant compte de l'expérience vécue des individus, de leurs perceptions subjectives ainsi que de la nature de leur relation avec les personnes avec lesquelles ils interagissent en ligne. Ces informations ne sont pas disponibles dans la simple observation des traces en ligne. Inversement, nous ne pouvions pas non plus nous appuyer uniquement sur des entretiens, puisque nous voulions pouvoir analyser des phénomènes dont les individus pouvaient ne pas se souvenir ou qu'ils pouvaient ne pas considérer comme pertinents.

C'est pour répondre à ce double besoin que nous avons conçu une technique innovante permettant de réaliser simultanément, d'une part, une observation directe – peu filtrée par l'individu – des pratiques effectives associées à Facebook et, d'autre part, un examen du contexte social et du sens conféré par les individus à ces pratiques numériques.

La collecte

Après la portion plus classique de l'entretien, la « visite commentée » était réalisée à l'aide de l'historique d'activités du profil de l'individu sur Facebook. En plus de ses publications ou « statuts », nous pouvions ainsi documenter plusieurs autres types de traces de son activité sur la plate-

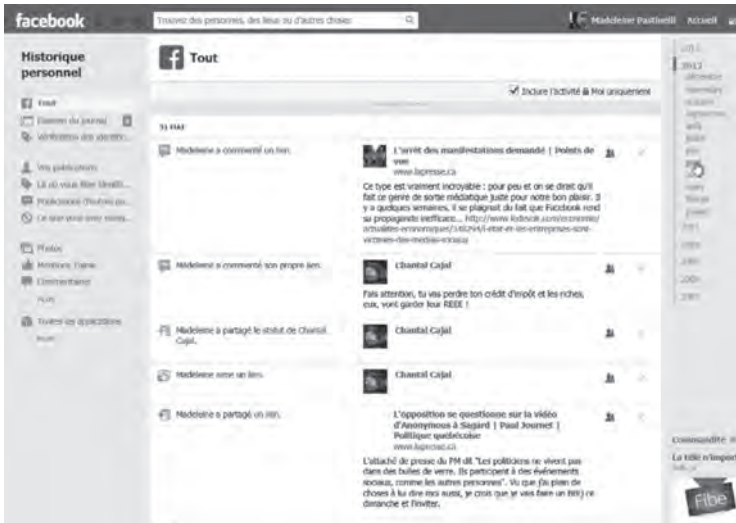
4. Ce projet fut réalisé à la demande du Centre d'études sur les médias (CEM).

forme: ses commentaires sur les publications de tiers, ses signaux de réaction tels que le « j'aime », etc.

Comme nous ne pouvions pas aborder avec les personnes rencontrées l'entièreté de leur historique Facebook sur une période de plusieurs mois, et afin de contenir l'analyse dans des paramètres comparables, nous avons sélectionné une série de huit à dix courtes périodes de deux jours à des dates spécifiques (avant, pendant et après le Printemps érable) à l'intérieur desquelles nous avons examiné l'historique des activités pour chaque personne⁵. La figure 11.2 illustre la ligne du temps à partir de laquelle, durant la visite commentée, nous pouvions atteindre les dates choisies.

FIGURE 11.2

Historique Facebook avec la liste des dates (capture d'une vidéo Camtasia)



Une fois que l'individu avait saisi son identifiant et son mot de passe, c'est le plus souvent une personne de l'équipe de recherche qui contrôlait l'ordinateur et qui naviguait vers les dates sélectionnées, ouvrant

5. Dans le cas où aucune trace d'activité n'était présente à une période donnée, l'intervue la remplaçait par les deux jours « actifs » les plus rapprochés de la date de référence, jusqu'à concurrence de la date de référence suivante. En outre, à sa discrétion, l'enquêteur pouvait renoncer à la visite de l'historique pour deux périodes, dans les cas où l'entretien aurait été beaucoup trop long en raison d'un historique particulièrement riche.

systématiquement toutes les traces dans un onglet à part, afin que cette phase de l'entretien se déroule de façon fluide. Pour chaque entrée, l'on demandait à l'individu d'expliquer à quoi correspondait l'information échangée, qui étaient les gens dont l'individu commentait l'activité ou qui commentaient la sienne, la nature de son lien avec eux, ce qu'il pensait de leurs échanges, etc. En outre, pour chacune des périodes, après avoir fait le tour des traces, l'on demandait également ce qui se passait de particulier à ce moment-là dans la vie de l'individu et dans son cercle d'amis.

Avec cet exercice, nous souhaitions notamment documenter (par l'observation directe du contenu à l'écran, mais aussi à l'aide du discours des individus au sujet de leur expérience) la façon dont les jeunes interagissaient avec leur entourage. Nous cherchions ainsi à connaître le type de discussions qu'ils avaient pu avoir durant la crise, en ligne et hors-ligne, avec la part de leur réseau social qui est en ligne, mais aussi avec celle qui ne l'est pas.

L'analyse

Alors que la portion classique des entretiens a donné naissance à des matériaux empiriques analysés de manière assez conventionnelle, la « visite commentée » a pour sa part fait l'objet de divers modes d'analyse spécifiques. Chaque membre de l'équipe de recherche a développé sa propre démarche, ce qui nous permet d'illustrer ici la variabilité et la fécondité analytique de cette technique. À quelques reprises, en cours d'analyse, nous avons croisé nos résultats préliminaires afin de nous alimenter mutuellement. Ces sessions de travail transversales ont permis d'enrichir les analyses produites par chaque méthode.

Premièrement, dans une approche descriptive et comparative des traces, il est possible de structurer l'analyse en s'appuyant sur la réalisation d'une base de données présentant un inventaire complet de chaque trace enregistrée (période, type de trace, émetteur initial du contenu, lien avec la grève, moment de la discussion autour de la trace dans l'enregistrement vidéo, etc.). Dans notre projet, cette base de données permettait d'abord de décrire le corpus dans son ensemble et ainsi de situer la place des traces relatives au Printemps érable parmi les autres (voir figure 11.3).

Ensuite, une sélection diversifiée de traces relatives au Printemps érable a fait l'objet d'un examen plus approfondi avec un retour transversal

FIGURE 11.3

Aperçu de la base de données des traces recueillies dans les vidéos Camtasia

	A	B	C	D	E	F	G
	Répétant	#	Période	Début	Type de trace	Émetteur	Porte sur les gèves?
2	OG#10	1	nov-12	00.01.51	aime un lien	autre	non
3	OG#10	2	nov-12	00.02.07	commente un lien	autre	non
4	OG#10	3	nov-12	00.02.41	écrit dans le journal de	répondante	non
5	OG#10	4	nov-12	00.03.46	aime un statut	autre	non
6	OG#10	5	oct-12	00.04.28	écrit dans le journal de	répondante	non
7	OG#10	6	oct-12	00.04.37	amis	/	non
8	OG#10	7	oct-12	00.05.35	aime un statut	autre	non
9	OG#10	8	oct-12	00.06.40	commente un statut	autre	non
10	OG#10	9	oct-12	00.06.48	commente une photo	autre	non
11	OG#10	10	oct-12	00.07.29	aime un lien	autre	non
12	OG#10	11	oct-12	00.08.30	commente une photo	autre	non
13	OG#10	12	oct-12	00.08.45	commente un lien	autre	non
14	OG#10	13	juin-12	00.09.35	aime un lien	autre	non
15	OG#10	14	juin-12	00.09.58	aime un lien	autre	oui
16	OG#10	15	juin-12	00.10.25	partage une photo	autre	non
17	OG#10	16	juin-12	00.10.46	aime un lien	autre	non
18	OG#10	17	juin-12	00.11.35	partage une photo	répondante	non
19	OG#10	18	juin-12	00.11.40	commente son propre	répondante	non
20	OG#10	19	28-mai-12	00.14.23	aime l'activité	autre	oui
21	OG#10	20	27-mai-12	00.15.03	partage une photo	autre	oui

sur les enregistrements vidéo. Cela a permis de générer une typologie des utilisations de l'information et des fonctions associées au contenu médiatique traditionnel présent sur la plateforme, tout en les mettant en relation avec les caractéristiques des individus révélées dans les entretiens, notamment en ce qui a trait à leur expérience vécue de la grève étudiante. Il s'agissait de mettre à profit des informations contextuelles qu'un logiciel automatisé d'extraction de traces n'aurait pas pu considérer.

Le visionnement systématique des captures vidéo a pour sa part donné lieu à deux types d'analyse. Une première stratégie a consisté à recourir au logiciel de prise de note Evernote pour soutenir une lecture attentive des fils de discussion (en particulier les chaînes de commentaires) et l'analyse du déroulement des débats entourant les traces portant sur le Printemps érable. Des images fixes tirées de copies d'écran des traces visuelles et textuelles utiles à l'analyse étaient colligées dans la fiche de l'individu concerné. Le logiciel effectue automatiquement une reconnaissance de caractères du texte dans les images, permettant une recherche par mots clés même dans les extraits de commentaires saisis en copie d'écran. Le dispositif a permis d'observer les formes que prenaient les échanges tout en tenant compte de la perspective de l'individu verbalisée durant la visite commentée (voir figure 11.4). Grâce à elle, nous avons pu obtenir des renseignements aussi bien sur le vécu de l'individu relativement

FIGURE 11.4

Exemple de trace consignée et annotée dans Evernote

14 JUIN 2012: A partagé et liké un mème sur la grève étudiante. "J'ai pas commenté, mais j'ai commenté en vrai avec mes amis et non sur la page" [ca. 10:20]



à la discussion analysée que sur les rapports entretenus hors-ligne par les divers intervenants.

Le dernier mode d'analyse, lui aussi lié directement au visionnement méthodique des captures vidéo, repose sur une démarche de type ethnographique. Il s'agissait de réaliser, pour chacun des individus rencontrés, une analyse des représentations et logiques qui organisent l'utilisation de la plateforme de façon générale et dans le contexte particulier de la grève. Cette première analyse, centrée sur la cohérence d'ensemble de chaque entrevue individuelle, a permis l'identification d'une série de pistes et de questions, lesquelles ont dans un second temps mené à aborder le corpus de façon transversale. Une nouvelle série de visionnements a soutenu l'exploration additionnelle de pistes et questions préalablement établies. Cette analyse a permis d'éclairer différentes dimensions de la circulation de l'information et de l'usage de Facebook, notamment quant aux différentes manières observées chez les individus de concevoir et d'utiliser la plateforme (comme un espace plus ou moins privé et centré ou non sur les échanges personnels), auxquelles correspondent des points de vue contrastés sur les usages qu'en font les autres.

Le projet Pratiques informationnelles sur l'emploi

La technique de la visite commentée a été utilisée plus tard dans un second projet. L'enquête portait sur les pratiques informationnelles – hors-ligne autant qu'en ligne – de personnes en recherche d'insertion en emploi (jeunes et nouveaux arrivants) (Gallant *et al.*, 2017). Ici, la portion des

entretiens consacrée à la visite commentée s'inscrivait au sein d'un ensemble plus large de techniques de collecte employées en entretien (incluant un calendrier âge-événement rétrospectif pour documenter les parcours et un générateur de noms pour documenter le réseau social).

La collecte

Pour ce projet, la « visite commentée » portait sur tous les sites Internet mentionnés lors de la première partie de l'entrevue sur les pratiques de recherche d'emploi. Ainsi, en dépit de certaines récurrences, nous ne pouvions pas toujours prévoir quels sites seraient visités. Pour chaque site, la visite commentée servait autant à obtenir l'impression des individus (« Qu'est-ce que vous pensez de l'information que vous trouvez sur ce site ? ») qu'à documenter de manière plus précise la façon dont il était utilisé (« Montrez-moi ce que vous faites en arrivant sur ce site »). Pour ce projet, l'observation se faisait uniquement sur les traces apparentes les plus récentes, plutôt que sur des périodes temporelles antérieures prédéterminées.

Par ailleurs, outre les sites eux-mêmes, nous observions aussi la façon dont l'individu naviguait sur le Web en général, puisqu'il contrôlait lui-même l'ordinateur pendant cette partie. Ensuite, nous l'amenions à visiter certains sites que nous désirions documenter systématiquement, et ce, même lorsque l'individu ne les avait pas mentionnés ou n'y avait pas de compte personnel, voire ne les connaissait pas.

L'analyse

L'analyse des pratiques numériques associées à la recherche d'emploi reposait sur trois sources complémentaires : ce qui en était dit dans la portion classique des entretiens, la visite commentée, ainsi qu'une observation documentaire indépendante concernant tous les sites Web mentionnés.

Pour commencer, un inventaire exhaustif de tous les sites Internet mentionnés en entrevue a été effectué. Un premier visionnement de l'ensemble des enregistrements des visites commentées a offert une première impression de leurs diverses utilisations. En complément, nous avons visité chaque site Web directement en ligne – et non uniquement dans les captures vidéo. Pour systématiser l'analyse, chaque site faisait

l'objet d'une fiche descriptive, rassemblant des informations selon diverses catégories classées de manière inductive : types de site, formes de modération ou contrôle, population ciblée, types de contenus informationnels, modes d'accès, apparence, etc. Le contenu de ces fiches descriptives était inséré dans un long tableau transversal, où figuraient aussi le code des individus et leurs usages du site (voir figure 11.5). Ce tableau contenait souvent des captures d'écran, mais celles-ci provenaient généralement plutôt de notre observation directe en ligne que des visites commentées.

FIGURE 11.5

Aperçu du tableau d'analyse des sites visités

Sites Internet formels		
Spécialisés dans l'emploi		
Emploi-Québec/placement en ligne [27] I2A1, I0I, I5JN	A17mh	
	A12qh	Pour chercher de l'emploi, A12qh commence à regarder ce site 4-5 mois avant de partir pour voir les emplois, voir comment ça fonctionne au Québec.
	A18qh	Pour chercher de l'emploi dans son domaine, mais l'utilise très peu.
	A16qf	Info sur le marché du travail en pharmacie.
	A11qh	Pour chercher de l'emploi dans son domaine.
	A11mf	Elle apprécie que ce soit une source fiable.
	A114sf	L'interviewée parlait du site d'Emploi-Québec qu'elle décrivait de non conviviale puisque l'accès au mode de recherche était difficile, le visuel n'était pas attrayant et puis les offres d'emploi étaient rarement intéressantes.
	I12qh	Pour chercher de l'emploi dans son domaine, il a également consulté les sections informatives, il trouve que l'information est bien structurée, « surtout quand on vient de l'étranger ».
	A19mf	Alerte venant de son courriel.
	A11qh	Utilise principalement les sources d'informations en ligne pour trouver de l'information sur l'emploi ou trouver de l'emploi.
Jobboom [14] 6A1, 4I, 4IN	A116mf	Trouver pleins de petites jobs, pas spécifique à son domaine.
	A15qf	Alerte venant de son courriel.
	I12qh	Emploi dans son domaine.
	I16qf	Peu utilisé, semble être plus pour les grands centres et peu pour les régions.
	J18mh	Pour trouver de l'emploi.
	J18mf	L'a consulté avant d'arriver au Québec, n'y est pas retourné, Emploi dans son domaine.
	J18mh	Dans son domaine, et quand dans le besoin un peu tout.

Enfin, les catégories d'informations de ce tableau transversal et des fiches descriptives des sites ont servi à construire une typologie des sites et de leurs usages. Ces informations pouvaient également être croisées avec les profils des individus pour dresser un portrait de leurs pratiques informationnelles en ligne.

En somme, une grande partie de l'analyse des sites Web s'est faite non pas sur les traces elles-mêmes, telles qu'elles sont visibles dans la capture vidéo de la visite commentée, mais plutôt en naviguant nous-mêmes directement sur les sites et en nous servant des enregistrements de la visite commentée tant en amont qu'en aval de ces observations indépendantes. Bien que nous ayons abordé l'utilisation de sites Internet durant la partie classique des entretiens, la visite commentée donnait l'occasion de constater *in situ* les raisons qui menaient les individus à utiliser un site

en particulier et la façon effective dont ils l'utilisaient. Les individus pouvaient aussi nous montrer directement les forces et les inconvénients de chaque site, dans la mesure où le fait d'avoir ces sites sous les yeux leur permettait de mieux se les remémorer. L'entretien classique n'aurait pas permis à lui seul de générer un corpus de données aussi riche.

Ces deux exemples de recours à la visite commentée montrent qu'il s'agit d'une technique très malléable. Elle permet soit d'étudier des modalités d'utilisation d'une plateforme spécifique ou d'une pratique numérique précise (auquel cas la visite commentée se trouve au cœur de l'analyse et la portion « classique » des entretiens est seulement complémentaire), soit de nourrir une analyse documentaire portant sur une diversité de sites Web orientés dans des pratiques sociales dont seulement une partie se déploie en ligne (alors c'est la visite commentée qui sert à enrichir une analyse dont le cœur se trouve ailleurs).

L'évaluation de la technique de la visite commentée

En prenant ces deux illustrations pour référence, nous proposons dans les prochaines pages un bilan critique des principales forces et limites de cette méthode.

Les forces principales

La force principale de la visite commentée réside dans la possibilité de bénéficier des explications contextuelles de l'individu et, surtout, de sa réflexivité sur le sens de ses pratiques. Ces renseignements supplémentaires concernent non seulement les traces visibles à l'écran, mais témoignent aussi de ce que vivait l'individu au moment où elles ont été produites.

Contextualisation et réflexivité

Chaque trace peut ainsi être située dans son contexte: qui sont les personnes avec lesquelles l'individu interagit? Quelle est la nature de ses relations avec les personnes qui apparaissent ou commentent le contenu des traces? Que se produisait-il dans la relation à ce moment? Les contenus se rapportent-ils à des choses qui étaient en train de se passer au travail ou

à l'école? Le croisement de ces informations nous a permis, par exemple, de montrer comment plusieurs jeunes pouvaient être très engagés hors-ligne dans la grève étudiante sans qu'il en paraisse rien sur leurs traces Facebook même aux temps forts de la crise.

La visite commentée permet également de documenter le sens et les motifs des actions effectuées en ligne, au lieu de se contenter d'une observation des comportements (voire d'inférences déductives possiblement douteuses – plutôt qu'empiriques – sur leur signification) faisant abstraction de la subjectivité des acteurs. Cette contextualisation et cette réflexivité ajoutent des éléments d'information que seul l'individu peut fournir.

L'observation rétrospective de pratiques effectives

Par ailleurs, la visite commentée opère une sorte de croisement entre l'observation documentaire et l'observation directe. En effet, cette technique ne dépend pas des souvenirs ou de la mémoire des individus et permet l'observation directe et rétrospective de pratiques effectives. Cet avantage s'avère particulièrement utile pour l'étude d'un phénomène ou de pratiques susceptibles de se transformer dans le temps.

La visite commentée permet en effet d'observer directement les pratiques des individus en contournant le filtre de leur conscience de celles-ci et celui de leur compréhension ou de leur interprétation de nos interrogations, problèmes qui se posent souvent en entretien. D'ailleurs, la consultation des traces permet de relever un certain décalage entre les pratiques rapportées en entrevue et celles effectivement visibles dans les traces. Ainsi, tandis qu'une personne affirmait passer peu de temps sur Facebook, la visite guidée de son historique d'activités nous a indiqué qu'elle consultait assez régulièrement le fil d'actualité puisqu'on retrouvait de nombreux « j'aime » dans l'historique⁶. De même, on observe des écarts entre les discours clamant un contrôle assez strict de l'exposition de la vie privée sur la plateforme et la gestion effective de la liste d'amis et des paramètres de confidentialité.

Le projet sur l'utilisation de Facebook pendant la grève étudiante met particulièrement en lumière une autre grande force de la méthode, liée

6. Il se peut que, pour certains individus du moins, ces « clics » soient concentrés sur une courte période. Une analyse plus fine de la densité temporelle des traces permettrait de le dire.

au fait que l'observation se réalise rétrospectivement. La recherche n'interfère pas avec les actions enregistrées et analysées, ce qui permet d'éviter que les individus ne modifient leur comportement parce qu'ils se savent observés.

En outre, l'observation rétrospective donne le loisir d'effectuer une étude rigoureuse des contenus enregistrés pendant la visite. Grâce aux enregistrements vidéo, l'analyse ne repose pas uniquement sur des notes d'observation initiales et prises à chaud (et complétées par nos souvenirs), mais s'appuie sur l'ensemble des données, qui peuvent faire l'objet d'un examen en profondeur rétrospectif. Cet aspect de la visite commentée renforce l'analyse inductive, en permettant de remarquer des éléments qui pourraient être passés inaperçus lors d'une observation initiale. Les enregistrements favorisent donc un retour itératif beaucoup plus poussé que ce que permet la relecture des notes d'observation, lesquelles sont, par nature, esquissées et incomplètes.

Le caractère éthique de l'observation et de la collecte

Comme l'observation est rétrospective, la visite commentée présente aussi un avantage sur le plan de l'éthique de la recherche, en permettant d'obtenir *a posteriori* un consentement pour l'observation de pratiques antérieures, ce qui laisse aux individus la maîtrise de ce qu'ils acceptent de montrer. L'individu peut acquiescer à l'observation avant que celle-ci ne se réalise, mais après le déploiement effectif des activités étudiées. Puisque l'individu nous accompagne ensuite tout le long de l'observation elle-même, cette méthode offre de surcroît l'occasion d'un consentement éclairé continuellement renouvelé. Au lieu de donner carte blanche à l'équipe de recherche pour observer toutes les activités archivées, ce qui se produit inévitablement lorsque les traces sont capturées « en bloc », le déroulement de la visite commentée permet à l'individu de demeurer constamment conscient des informations qui sont récoltées. À tout instant, il pourrait interrompre la visite ou demander l'exclusion d'une trace du corpus si l'accès au contenu affiché à l'écran l'incommode. Ce droit de se retirer ou de sauter certaines traces était d'ailleurs rappelé aux individus au moment d'entreprendre cette phase de l'entretien.

Des données riches et denses codées « à la main »

Les dernières forces de ce dispositif concernent l'ampleur et la qualité du jeu de données constitué, favorisant des analyses qualitatives riches. Alors qu'une capture « en masse » rend presque obligatoire un traitement automatisé des traces, la collecte sélective de traces (par échantillonnage raisonné et sélection temporelle) réduit efficacement le volume de données et rend le corpus codable « à la main » en maximisant la proportion de données porteuses d'informations pertinentes en lien avec la problématique de la recherche.

De plus, nous l'avons vu, la visite commentée produit des données riches et « denses », puisqu'elles sont accompagnées de « couches » d'informations complémentaires fournies tant par leur observation « en contexte » que par le discours des acteurs dont les traces sont observées (Latzko-Toth, Bonneau et Millette, le présent ouvrage). Ce métadiscours sur les traces des pratiques numériques permet de leur attacher des éléments du contexte de leur production et du sens qu'elles revêtent pour les individus.

Les limites à considérer

Dans toute analyse d'une visite commentée, il faut néanmoins garder à l'esprit les angles morts de la technique. Certaines limites sont liées aux archives d'une plateforme. Ainsi, les paramètres techniques de l'historique d'activités de Facebook font que les traces archivées ne concernent que les contenus avec lesquels l'individu a activement interagi sur la plateforme, c'est-à-dire en les partageant, en les commentant ou en cliquant sur le bouton « j'aime ». Cela ne permet pas de connaître les autres contenus auxquels l'individu aurait pu être exposé (par exemple, par une lecture « passive » du fil d'actualité) sans y réagir, ni même les contenus médiatiques affichés par un tiers qu'il aurait consulté sur le site d'origine sans revenir à Facebook pour y réagir d'une manière qui « laisse des traces » dans son historique.

Par ailleurs, lorsque l'objectif de recherche consiste à documenter un éventail de sites, la technique se heurte à la capacité de l'individu à se remémorer *in situ* tous les sites qu'il aurait pu consulter. Même sa conscience des sites visités peut être restreinte, par exemple dans le cas de jeunes qui disent avoir consulté « Google » pour leur recherche d'em-

ploi, sans que nous soyons en mesure de savoir sur quel site leurs recherches les ont fait aboutir en pratique. La capacité de prétendre à une certaine exhaustivité est ainsi restreinte.

* * *

La visite commentée est une technique de collecte de données denses à petite échelle. Elle repose sur une exploration et une saisie sélectives et commentées des traces d'activité en provenance des médias sociaux et d'Internet. Ces traces sont observées dans leur contexte d'origine et en présence de la personne qui les a produites. L'interaction avec un membre de l'équipe de recherche, en même temps que l'observation initiale, construit un matériel consistant qui enrichit l'analyse et apporte une épaisseur supplémentaire aux données de traces en leur associant diverses couches de contexte, de description et de sens. Cette technique comporte aussi une étape d'entretiens semi-dirigés plus classiques, qui permettent d'associer des profils de vécus hors-ligne à l'analyse des pratiques numériques observées durant la visite commentée. Enfin, en dépit de certaines limites, la méthode possède aussi une grande adaptabilité à différentes problématiques de recherche pour lesquelles elle peut être plus ou moins centrale.

Cette technique produit des données riches sur les traces d'activités effectives, qui sont « densifiées », d'une part, par les explications contextuelles et la réflexivité des sujets sur leurs pratiques et, d'autre part, par des données discursives récoltées en entretien sur les profils sociodémographiques et les parcours hors-ligne. L'analyse de ces données nuancées permet d'éviter des généralisations réductrices découlant d'un regard homogénéisant sur des profils d'individus (par exemple, sur les jeunes, trop souvent groupés sous le vocable trompeur de « natifs du numérique »). En effet, la visite commentée favorise une analyse riche permettant d'observer la diversité des usages et des compétences.

Références

BEAUVISAGE, Thomas, « Compter, mesurer et observer les usages du Web : outils et méthodes », dans Christine BARATS (dir.), *Manuel d'analyse du Web en sciences humaines et sociales*, 2^e édition, Armand Colin, 2016, p. 199-222.

- BOYD, danah, Scott GOLDBER et Gilad LOTAN, « Tweet, tweet, retweet: Conversational aspects of retweeting on Twitter », dans *HICSS, 10 Proceedings of the 2010 43rd Hawaii International Conference on System Sciences*, IEEE Computer Society, 2010, p. 1-10.
- BRUNS, Axel et Jean BURGESS, « Researching news discussion on Twitter: New methodologies », *Journalism Studies*, vol. 13, n^{os} 5-6, 2012, p. 801-814.
- DUBOIS, Elizabeth et Heather FORD, « Trace interviews: An actor-centered approach », *International Journal of Communication*, vol. 9, n^o 25, 2015, p. 2067-2091.
- GALLANT, Nicole et al., *Pratiques informationnelles en matière d'insertion en emploi. Le cas des jeunes et des immigrants récents au Québec*, rapport de recherche commandité par le ministère du Travail, de l'Emploi et de la Solidarité sociale et lui étant remis, 2017.
- GALLANT, Nicole, Guillaume LATZKO-TOTH et Madeleine PASTINELLI, *Circulation de l'information sur les médias sociaux pendant la grève étudiante de 2012 au Québec*, Centre d'études sur les médias, 2015.
- HONG, Lichan, Gregorio CONVERTINO et Ed H. CHI, « Language matters in Twitter: A large scale study characterizing the top languages in Twitter characterizing differences across languages including URLs and hashtags », dans *Proceedings of the Fifth International AAAI Conference on Weblogs and Social Media (ICWSM '11)*, Menlo Park, Californie, 2011, p. 518-521.
- MCVEIGH-SCHULTZ, Joshua et Nancy K. BAYM, « Thinking of you: Vernacular affordance in the context of the microsocial relationship app, couple », *Social Media + Society*, vol. 1, n^o 2, 2015, p. 1-13.
- MILLETTE, Mélanie, « Contre-publics et médias sociaux: le cas du collectif francophone taGueule! au Canada anglais », dans Serge PROULX, José Luis GARCIA et Lorna HEATON (dir.), *La contribution en ligne: pratiques participatives à l'ère du capitalisme informationnel*, Presses de l'Université du Québec, Québec, 2014, p. 91-102.
- MILLETTE, Mélanie, Josianne MILLETTE et Serge PROULX, « Hashtags et casseroles: de l'auto-organisation du mouvement social étudiant », *WI Journal of Mobile Media: Out of The Mouths of the « casseroles »*, vol. 6, n^o 2, 2012.
- MOLLER JORGENSEN, Kristian, « The media go-along: Researching mobilities with media at hand », *MedieKultur: Journal of Media & Communication Research*, vol. 32, n^o 60, 2016, p. 32-48.
- ROBARDS, Brady et Siân LINCOLN, « Uncovering longitudinal life narratives: Scrolling back on Facebook », *Qualitative Research*, vol. 17, n^o 6, 2017, p. 715-730.

CHAPITRE 12

L'entretien d'autoconfrontation : analyse d'une activité de navigation sur Internet

Daniel Pélissier

De nombreuses études portent sur les traces numériques dites volontaires, comme les commentaires des réseaux socionumériques ou encore les messages de forum. Les possibilités de collecte automatique des données ont aussi favorisé la réalisation d'analyses massives de *tweets* ou de commentaires Facebook. Malgré cela, l'internaute qui navigue et ne laisse ni message, ni commentaire, ni aucune trace volontaire est peu étudié. Certaines approches portent sur les traces de navigation (Gong *et al.*, 2015; Bastard *et al.*, 2017) enregistrées sur les serveurs Internet et mettent en évidence certains comportements traçables des internautes. Cependant, la boîte noire des mécanismes cognitifs est rarement ouverte (Prost *et al* 2013).

Certaines études ont cherché à évaluer la part des internautes qui naviguent sur Internet, ce comportement étant souvent considéré comme passif. La répartition des internautes entre actifs et passifs évoluerait alors vers plus de passivité (IFOP, 2013), observation qui a été érigée en principe fondé sur l'intuition, selon la règle des 90-9-1¹ : 90 % des internautes seraient des badauds (*lurkers*), 9 % des contributeurs intermittents et 1 % des contributeurs réguliers. Plus précisément et scientifiquement, Bastard

1. Principe proposé par Nielsen (2006).

et al. (2017) repèrent 17 % d'internautes usagers de Facebook non actifs² qui regardent sans publier et utilisent ce dispositif comme une fenêtre ouverte sur le monde.

Au-delà de ces données sensiblement différentes, la navigation ne peut être considérée comme un processus passif ou une activité réflexe. L'internaute réalise des choix, peut suivre une perspective, prend en compte un contexte, etc. Nous partageons ainsi la définition de la navigation comme « exercice de construction des espaces subjectifs de compréhension » (Ghitalla et Lenay, 2002, p. 12). La navigation, perçue comme une construction de sens, exige alors des méthodes de nature qualitative permettant de comprendre les motivations des internautes que les données massives laissent simplement imaginer.

Ce chapitre analyse les apports d'une méthode qualitative, l'autoconfrontation simple, qui permet d'expliquer la navigation des internautes. Nous fondons cette présentation sur les résultats d'une recherche³ concernant l'exploration de sites Internet d'entreprises par de jeunes personnes diplômées de l'enseignement supérieur dans le cadre d'une candidature à un emploi. Plus précisément, nous présentons notre méthode en la replaçant dans une perspective historique et disciplinaire en appui à quelques résultats significatifs issus de la recherche, afin d'illustrer le potentiel de cette approche. Cela nous permet d'envisager les apports de l'autoconfrontation simple, notamment la possibilité d'accéder à la conscience de l'internaute navigateur avant de conclure sur les limites et perspectives de cette méthode.

La méthode de l'autoconfrontation simple

Notre recherche s'est intéressée à la réception des données issues de différents sites Internet utilisables pour une recherche d'emploi par de jeunes

2. Ce pourcentage, plus faible que ceux de l'IFOP, met plus en évidence, selon nous, une difficulté à cerner de façon homogène la notion d'activité qu'une évolution marquée vers un comportement numérique plus participatif. Les différences d'échantillon et de méthode rendent ces études non comparables.

3. Recherche réalisée dans le cadre d'un doctorat soutenu le 20 juin 2018 sous la direction de Martine Corral Regourd, IDETCOM, et Robert Boure, université de Toulouse, et accompagné par Laurence Leveneur, MCF IDETCOM, IUT de Rodez, et Isabelle Vidalenc, MCF IDETCOM, IUT de Rodez. Nous remercions également Nicole Boubée (ESPE), Patrick Sabi et Annie Fraison, École d'ingénieur 3IL, ainsi que Laetitia du Chaffaut et Sarah François, Toulouse School of Management, pour leur aide.

personnes diplômées (site institutionnel d'une entreprise, page Facebook sur le recrutement et plateforme de témoignages de personnes salariées de cette même organisation⁴). La question de recherche portait ainsi sur la façon dont ces jeunes personnes diplômées reçoivent ces informations lorsqu'elles préparent un entretien de recrutement ou une candidature spontanée. La présence numérique des entreprises étant très variée, nous avons centré l'observation sur une mise en situation concernant le secteur bancaire, très impliqué dans la communication numérique de recrutement. Pour comprendre cette activité de navigation de jeunes personnes diplômées, nous nous sommes alors penchés sur la méthode de l'autoconfrontation, dont nous présenterons les principes avant de préciser son adaptation aux usages numériques d'Internet.

Historique et ancrage de la méthode

Dans un entretien d'autoconfrontation, « chaque sujet concerné est confronté à des traces filmées de son activité, qu'il commente » (Clot et Leplat, 2005, p. 307). Une personne participante commente librement la trace de son activité qui est captée sous forme de vidéo, la plupart du temps; il s'agit ensuite d'analyser ce discours ainsi que le film de l'activité. Les premières mentions de cette méthode remonteraient aux travaux de Von Cranach au début des années 1980 (Theureau, 2002) dans le cadre d'une théorie sociale sur la compréhension des ressorts de l'activité humaine. Les applications de cette méthode ont porté en particulier sur l'étude de situations de travail (Mollo et Falzon, 2004). L'activité est alors celle d'une personne salariée qui commentera une vidéo de son travail. Le domaine de l'éducation (Leblanc *et al.*, 2013) a aussi choisi cette approche pour comprendre et aider les personnes enseignantes dans leurs pratiques réelles en classe. Enfin, le sport (Boyer *et al.*, 2015) est également intéressé par cette démarche, qui lui permet de visualiser et d'améliorer notamment les performances.

Il existe deux types d'autoconfrontation, simple et croisée. La simple consiste à faire commenter l'activité par la personne participante seulement. L'adjectif « simple » ne s'attache ni à une facilité de mise en œuvre

4. Ces sites ont été choisis parce qu'ils étaient représentatifs de ceux utilisés pour la recherche d'emploi et portaient sur le problème de la construction identitaire de l'entreprise dans un contexte de recrutement que nous souhaitons étudier.

ni à un manque de profondeur des données. Il s'oppose plutôt à l'autoconfrontation croisée consistant à faire commenter l'activité par une autre personne. La confrontation de la personne participante à sa propre activité est alors croisée avec les analyses d'un tiers. Cette dernière approche s'inscrit plus dans une visée de transformation sociale (Clot et Leplat, 2005), d'apprentissage, adaptée notamment à des situations d'intervention du chercheur sur un terrain. Dans notre recherche, la finalité était plus compréhensive et a privilégié une autoconfrontation dite simple.

Cette méthode s'inscrit dans l'ensemble plus large des méthodes de verbalisation (Forget, 2013). Elle se distingue de deux approches qualitatives voisines : les enquêtes effectuées à partir de discours et la méthode *think aloud* (parfois traduit par « penser tout haut »). Une analyse de discours décontextualisés vaut notamment par sa distance par rapport à l'expérience étudiée, car elle « saisit la représentation articulée à son contexte expérientiel et l'inscrit dans un réseau de signification » (Blanchet et Gotman, 2007, p. 25). L'éloignement du contexte de l'expérience tendrait à favoriser des discours généralisants et finalement hors contexte, mal situés. Par exemple, pour notre recherche, une collecte de données fondée sur des questions sur le recrutement, fussent-elles ouvertes, fournirait davantage des discours construits sur la situation que des données proches de la conscience des personnes participantes. Un entretien d'explication après l'activité pourrait avoir les mêmes limites (Forget, 2013). Certains peuvent préférer pour leur problématique ce type d'approche, mais nous l'avons écartée dans cette recherche, en considérant que la réduction de l'écart entre le discours et la situation de communication étudiée favorisait des interprétations plus proches des mécanismes cognitifs des personnes participantes.

La méthode *think aloud* s'appuie sur une situation concrète de navigation qui évite les dérives de discours déclaratifs parfois éloignés des pratiques réelles. Elle consiste à demander à la personne participante de commenter son activité pendant qu'elle se déroule (Ericsson et Simon, 1993), de façon concomitante et non rétrospective comme pour l'autoconfrontation. Pinsky et Theureau notaient par exemple que « l'autoconfrontation, contrairement aux autres formes de verbalisation [...], ne perturbe en rien la dynamique de l'action en cours » (1992, p. 35). Par ailleurs, si la méthode *think aloud* est possible dans certaines situations, certains contextes empêchent sa mise en œuvre, par exemple dans le cas où l'activité utilise la voix ou si le commentaire gêne l'activité (Rix-Lièvre, 2010,

p. 374). L'autoconfrontation permet ainsi, en comparaison avec la méthode *think aloud*, d'observer une activité naturelle sans les interruptions liées aux commentaires et d'analyser ainsi plusieurs sources de données, le film de l'activité et son commentaire.

L'usage de cette méthode pour la compréhension d'une navigation Internet a nécessité quelques ajustements.

L'autoconfrontation simple

Appliquée à Internet, l'autoconfrontation simple consiste à enregistrer l'activité de la personne participante, sa navigation Internet dans ce cas, puis à la confronter à cette trace filmée lors d'un entretien (Boubée, 2010).

Notre pratique de recherche a été la suivante :

- enregistrement sous forme de vidéos (logiciel OBS studio⁵) de la navigation des jeunes personnes diplômées sur trois sites Internet (site carrière de la BNP⁶, page de recrutement Facebook⁷, témoignages de personnes salariées⁸) ;
- explication de la navigation par la personne jeune diplômée qui commente l'enregistrement de son parcours⁹ ;
- échange informel sur l'entretien d'autoconfrontation.

Chaque entretien dure ainsi environ une heure trente pour l'ensemble des phases et forme deux corpus : une vidéo de l'activité de navigation pour avoir une idée du processus complet, et les commentaires des personnes participantes sur ces enregistrements.

Dans notre recherche, notre échantillon est constitué de 21 jeunes personnes diplômées du programme de l'enseignement supérieur français de niveau bac+5, en deuxième cycle universitaire. Elles sont volontaires pour participer et sont issues de deux institutions de formation différentes : une école d'ingénieur informatique (sigle 3IL) et un institut d'administration des entreprises (IAE).

5. Voir obsproject.com/.

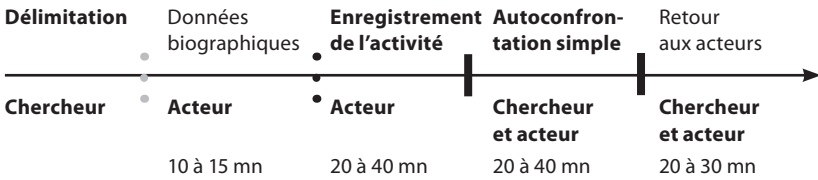
6. Voir group.bnpparibas/emploi-carriere.

7. Voir fr-fr.facebook.com/bnpparibascareers/.

8. Voir choosemycompany.com/fr/ et www.glassdoor.fr/.

9. Un exemple de vidéo est présenté sur presnumorg.hypotheses.org/235.

FIGURE 12.1

Mise en œuvre de l'autoconfrontation simple

La mise en œuvre de cette méthode a exigé que l'on fasse de nombreux choix pour l'adapter au contexte numérique. Nous avons en particulier travaillé sur la partie du retour aux personnes participantes afin d'améliorer la validité des données collectées et le traitement du double corpus en liant activité et commentaires (Pélissier, 2017).

Un des apports de cette méthode a trait aux données qu'elle permet d'analyser. Nous illustrerons cela par les résultats de notre recherche permettant d'expliquer une activité de navigation sur des sites Internet.

L'analyse des motivations et des logiques de navigation : temps de connexion et activité réelle

Cette méthode permet d'étudier les trajectoires de navigation qui ne laissent pas de traces volontaires et qui représentent pourtant une part importante des activités en ligne. Nous présenterons deux types d'analyse des motivations d'une activité de navigation en donnant l'exemple de la lecture des témoignages de personnes salariées de l'organisation d'une part et des pauses pendant la navigation d'autre part.

Les statistiques des sites Internet fournissent des temps de connexion et ces durées sont surveillées étroitement par les gestionnaires de communauté (*community managers*) pour évaluer la réussite de leur site. On tient alors pour implicite que le temps de connexion est un temps de consultation.

Nos résultats montrent pourtant une réalité complexe. D'abord, la consultation n'a rien de neutre, de complète ou de linéaire, mais est orientée, bricolée. Ensuite, le temps d'affichage de la page peut aussi englober une absence de l'internaute.

La lecture bricolée des pages Internet

Plus finement que les statistiques de sites Internet, les procédés d'oculométrie (*eye tracking*) permettent de repérer les zones observées pendant une navigation. Plus la zone-écran est consultée, plus elle serait importante pour l'internaute. Si ces techniques permettent d'approcher des comportements, elles demeurent muettes sur leurs raisons profondes – le parcours est précisé, mais pas ses motivations.

Les entretiens d'autoconfrontation ont, quant à eux, permis de comprendre les pratiques mises en œuvre par ces jeunes personnes diplômées. Les commentaires précis sur l'activité de navigation ont favorisé la compréhension fine de certains usages. De Certeau a montré comment la réception des textes fait l'objet d'une lecture bricolée : « Les lecteurs sont des voyageurs ; ils circulent sur les terres d'autrui, nomades braconnant à travers les champs qu'ils n'ont pas écrits » (1990, p. 251). Les internautes rencontrés dans cette recherche ne se conforment pas forcément à l'influence des personnes qui ont conçu les dispositifs, mais bricolent, construisent un sens et des usages selon leurs représentations individuelles et sociales, décident d'un parcours aux motivations variées.

Nous illustrons ce que cette méthode permet justement de montrer en présentant l'observation commentée de la navigation des jeunes personnes diplômées sur un site de témoignages de personnes salariées. Ce type de site, indépendant des sites institutionnels des organisations, rassemble des avis qui peuvent être utilisés par les individus candidats au recrutement pour choisir l'organisation ou préparer un entretien (figure 12.2).

Loin de prendre ces commentaires au pied de la lettre, ces jeunes personnes diplômées réalisent des opérations complexes de filtrage, d'interprétation, de mise en contexte, etc. Ainsi, la longueur des témoignages est un indicateur de fiabilité et d'intérêt partagé par plusieurs. La longueur n'est pas le seul indice d'intérêt de l'avis. Sur un des sites observés, certains avis apparaissaient en double. Les raisons potentielles évoquées par ces internautes sont multiples : erreur de manipulation, volonté de promouvoir son avis, absence de modérateur, etc. À l'inverse de la longueur, les doublons décrédibilisent la qualité de la source. Ces jeunes personnes diplômées partagent ainsi une théorie de repérage des avis pertinents acquise par les usages antérieurs d'autres sites de même type. La longueur ou les doublons sont des opérateurs de tri efficaces pour

FIGURE 12.2

Glassdoor, exemple de témoignage de salariés

décèler les avis pertinents qui ont plus de chance d'être fiables. Dans un monde numérique traversé par les fausses informations (*fake news*), ces internautes ont mis en œuvre des tactiques adaptatives pour fiabiliser l'information, dans la mesure du possible.

L'ordre de lecture des avis est un autre bricolage qui a pu être décrypté par la méthode d'autoconfrontation. Plusieurs personnes participantes ont ainsi expliqué leurs usages de consultation des avis quand ceux-ci sont classés par jugement positif ou négatif. Elles ont souvent choisi un ordre de parcours. Leur logique de navigation n'est que partiellement influencée par le dispositif. Certaines commencent volontairement par des avis positifs, d'autres par les négatifs. Cette théorie personnelle s'est construite sur une autoanalyse et sur une habitude, pas forcément de ces dispositifs, mais de l'évaluation en général. Dans certains cas, le risque d'être « rebuté » par les avis négatifs l'emporte et les positifs sont alors consultés en premier. Dans d'autres, enfin, les personnes ne regardent que les négatifs, comme cette jeune diplômée qui résume sa position sans ambages : « tout ce qui est apprécié, ça sert à rien. Les bonnes critiques en général, c'est lié à la personne, pas au cas général ».

Ainsi, l'ordre de lecture des avis n'est pas partagé et dépend de choix personnels qui illustrent les marges de liberté que se donnent ces jeunes personnes diplômées par rapport à l'influence des dispositifs. Contrai-

rement aux théories élaborées sur les repères de lecture (longueur, doublons) qui paraissent plus générales, les règles d'ordre de lecture semblent plus individuelles. Cependant, la fréquence de l'apparition de ces commentaires montre un questionnement partagé qui renvoie à la pratique de ce type de site. La tendance à l'évaluation généralisée a contribué aussi à l'élaboration de procédés d'appropriation de ces notations. Ces jeunes personnes diplômées ont baigné dans cet univers évaluatif, que ce soit pour les achats ou les services, ou encore pendant leurs études. La lecture de ces internautes apparaît ainsi comme un procédé complexe que le temps de connexion enregistré dans les statistiques des sites Internet ou le repérage de zones observées par un dispositif d'oculométrie ne résument que très partiellement.

L'absence de lecture ou l'escale de navigation

Alors que les données massives ne permettent d'accéder qu'aux traces de comportements de navigation, la méthode de l'autoconfrontation a facilité la compréhension des raisons de certains phénomènes, mais a aussi évité des contresens. En effet, certains temps de connexion peuvent correspondre à des pauses dont les motivations sont très variables. Une page qui apparaît figée dans le film de la navigation peut être lue ou, au contraire, signifier une pause de l'internaute pendant sa navigation.

Le commentaire de l'internaute pendant la confrontation permet de donner du sens à ces périodes qu'un serveur Internet se limiterait à recueillir. La pause peut alors correspondre à une adaptation aux contraintes techniques du moment et ce bricolage technologique renvoie alors à l'ouverture d'onglets dans le navigateur. De même, la pause peut renvoyer à la réalisation d'une autre activité pas forcément liée à la lecture de la page : une réflexion, une discussion, l'usage du téléphone mobile, etc.

Ainsi, le temps de connexion enregistré par le serveur peut correspondre à des pauses dont les raisons sont très éloignées d'une lecture plus ou moins attentive de la page. Ces temps de pause sont comme des escales de navigation qui n'ont pas de sens pour un serveur de données, mais font partie intégrante de la navigation. Cette dernière est intégrée plus largement dans la variété des activités d'une vie humaine ; une statistique de connexion, si elle donne une idée d'un phénomène, peut aussi cacher cette dimension sociale de l'activité de l'internaute pourtant fondamentale.

Les apports et limites de l'autoconfrontation simple

Le principal apport intrinsèque de cette méthode est la compréhension des motivations de la navigation. En effet, les traces de navigation informent sur le comportement et peuvent donner l'illusion « d'une forme plus élevée d'intelligence et de connaissance » (traduction libre de Boyd et Crawford, 2012, p. 663). Les données massives collectent des traces de comportement, mais ne peuvent pas fournir d'analyse des motivations d'un individu dont le comportement complexe ne peut se résumer à quelques informations chiffrées (Rouvroy, 2014). Elles peuvent laisser supposer des corrélations dont la teneur est souvent sujette à caution. La quantité des données ne correspond pas à leur richesse ou à leur épaisseur (Ollion et Boelaert, 2015) et l'intérêt de la méthode qualitative présentée est d'obtenir des résultats dépassant les limites de ces données massives. L'autoconfrontation permet en effet d'accéder à la partie invisible de la situation observée (Cahour et Licoppe, 2010) qui ne laisse aucune trace susceptible d'être quantifiée. Si l'observation de l'activité permet une première approche des mécanismes cognitifs impliqués dans une navigation, elle ne favorise pas une compréhension profonde. Une verbalisation rétrospective assistée permet d'accéder à la conscience de la personne participante (Theureau, 2010). Cette approche de nature phénoménologique considère l'expérience comme la mise en œuvre d'une interaction entre un monde extérieur et une conscience située dont l'accès ne peut se faire que par l'action qui précède la représentation (Rix-Lièvre, 2010).

Un autre apport de cette méthode est de permettre au chercheur ou à la chercheuse de saisir une trajectoire, une dynamique. Le temps et la temporalité sont des dimensions essentielles de la compréhension des usages d'Internet (Domenget *et al.*, 2017). La compréhension de ce facteur historique est favorisée par le recul possible grâce aux données accumulées pendant la trentaine d'années de pratiques d'Internet. Cette approche temporelle longue permet de proposer des perspectives historiques (Flichy, 2001). La méthode de l'autoconfrontation simple s'intéresse plus au temps court d'une navigation Internet. Les deux peuvent être liées puisque les motivations de l'internaute expliquant sa navigation puisent dans l'historique de ses usages, dans ses représentations individuelles et sociales. Plus précisément, l'autoconfrontation a été l'occasion de mesurer les temps réels de parcours d'une page et de les associer aux motivations des internautes. Cette démarche décrit

alors la navigation comme un processus, une dynamique aux contours singuliers dont la finalité de construction de sens se déploie, non pas au détour d'un clic ou d'un *scroll* (défilement), mais dans une pensée complexe, itérative, mouvante et ancrée dans un historique d'usages.

Enfin, cette méthode présente des enjeux scientifiques variés axés sur les usages numériques. D'abord, cet effort scientifique évite des généralisations abusives qui continuent à diffuser des principes erronés. Par exemple, les *digital natives*¹⁰ (Prensky, 2001) sont un concept né d'une réflexion sans terrain qui enferme toute une génération dans un comportement homogène. Des recherches qualitatives (Cordier, 2015 ; boyd, 2016) ont démontré les limites de cette approche. De même, l'autoconfrontation, en dépassant les discours déclaratifs, peut compléter cette dernière orientation et apporter des nuances aux conceptions stéréotypées de certains analystes des données massives.

De plus, l'étude de la navigation fait écho à la réception médiatique dont nous avons analysé la filiation et les différences. Or, les analyses de réception des médias ont montré les jeux de pouvoir traversant les activités de réception et de production (Hall, 1994). Transférée au monde numérique, cette approche peut montrer comment l'internaute, face à l'influence des GAFAM, utilise des marges de liberté pour limiter l'emprise du pouvoir des géants d'Internet (Smyrnaioi, 2017). L'autoconfrontation favorise ainsi une analyse politique d'Internet qui se démarque d'une vision radicalement déterministe des dispositifs. Elle aborde plutôt l'internaute comme une personne active parfois consciente du jeu de pouvoir des dispositifs, qui joue, bricole, selon ses objectifs, ses connaissances, ses représentations.

Si cette mise en œuvre de l'autoconfrontation permet de mieux comprendre les motivations de certains phénomènes, elle présente aussi certaines limites. La mise en situation pose, notamment, la question de la naturalité du contexte. L'observation n'est jamais neutre pour la personne participante et peut modifier la pratique. Certains travaux recommandent une approche ethnographique (Lièvre et Rix-Lièvre, 2013) qui peut favoriser l'intégration du chercheur ou de la chercheuse. Même avec cette démarche proche du terrain, la question de la naturalité de l'activité reste posée. Notre recherche a soulevé des interrogations sur cette limite

10. Natifs du numérique.

de l'autoconfrontation, à laquelle il ne semble pas exister de réponse simple. La phase de retour aux personnes participantes après l'entretien, l'utilisation de sites Internet réels, l'attention aux questions portant sur ce thème pendant la mise en situation ont été, par exemple, des éléments de limitation de ce biais dans cette recherche. D'autres travaux (par exemple Gallant, Latzko-Toth et Pastinelli, 2015) ont analysé une sélection de l'historique de Facebook, ce qui peut poser certains problèmes de mémorisation, mais garantit une trace de navigation naturelle. Ni dispositif expérimental, ni situation naturelle, l'activité étudiée avec l'autoconfrontation se situe dans un entre-deux qui doit être interrogé.

Références

- BASTARD, Irène *et al.*, « Facebook, pour quoi faire ? Configurations d'activités et structures relationnelles », *Sociologie*, vol. 8, n° 1, 2017, p. 57-82.
- BLANCHET, Alain et Anne GOTMAN, *L'entretien*, Armand Colin, 2015.
- BOUBÉE, Nicole, « La méthode de l'autoconfrontation : une méthode bien adaptée à l'investigation de l'activité de recherche d'information ? », *Études de communication*, n° 35, 2010, p. 2-9.
- BOYD, danah, *C'est compliqué, les vies numériques des adolescents*, C&F Éditions, 2016.
- BOYD, danah et Kate CRAWFORD, « Critical questions for big data : Provocations for a cultural, technological and scholarly phenomenon », *Information, Communication & Society*, vol. 15, n° 5, 2012, p. 662-679.
- BOYER, Simon, Géraldine RIX-LIÈVRE et Michel RÉCOPÉ, « L'arbitrage de haut niveau, une affaire d'équipe », *Movement & Sport Sciences*, vol. 1, n° 87, 2015, p. 91-101.
- CAHOUR, Béatrice et Christian LICOPPE, « Confrontations aux traces de son activité », *Revue d'anthropologie des connaissances*, vol. 4, n° 2, 2010, p. 243-253.
- CLOT, Yves et Jacques LEPLAT, « La méthode clinique en ergonomie et en psychologie du travail », *Le travail humain*, vol. 68, n° 4, 2005, p. 289-316.
- CORDIER, Anne, *Grandir connectés, les adolescents et la recherche d'information*, C&F Éditions, 2015.
- DE CERTEAU, Michel, « Lire : un braconnage », dans *L'invention du quotidien, I, Arts de faire*, Gallimard, 1990, p. 239-255.
- DOMENGET, Jean-Claude, Bernard MIÈGE et Nicolas PÉLISSIER, *Temps et temporalités en information-communication, des concepts aux méthodes*, L'Harmattan, 2017.
- ERICSSON, Karl Anders et Herbert A. SIMON, *Protocol Analysis: Verbal Report as Data*, The MIT Press, 1993.
- FLICHY, Patrice, *L'imaginaire d'Internet*, La Découverte, 2001.
- FORGET, Marie-Hélène, « Le développement des méthodes de verbalisation de l'action : un apport certain à la recherche qualitative », *Recherches qualitatives*, vol. 32, n° 1, 2013, p. 57-80.

- GALLANT, Nicole, Guillaume LATZKO-TOTH et Madeleine PASTINELLI, *Circulation de l'information sur les médias sociaux pendant la grève étudiante de 2012 au Québec*, Centre d'études sur les médias, Sainte-Foy, 2015.
- GHITALLA, Franck et Charles LENAY, « Les territoires de l'information. Navigation et construction des espaces de compréhension sur le Web », *Les Cahiers du Numérique*, vol. 3, 2002, p. 51-63.
- GONG, Wei, Ee-Peng LIM et Feida ZHU, « Characterizing silent users in social media communities », dans *Proceedings of the Ninth International AAAI Conference on Web and Social Media*, 2015, p. 1-10.
- HALL, Stuart, « Codage/décodage », *Réseaux*, vol. 12, n° 68, 1994, p. 27-39.
- IFOP, *Observatoire des réseaux sociaux 2013*, Institut français d'opinion publique, Paris, 2013.
- LEBLANC, Serge, Luc RIA et Philippe VEYRUNES, « Vidéo et analyse *in situ* des situations d'enseignement et de formation dans le programme du cours d'action », dans Laurent VEILLARD et Andrée TIBERGHEN (dir.), *ViSA : Instrumentation de la recherche en éducation*, Éditions de la Maison des sciences de l'homme, 2013, p. 63-94.
- LIÈVRE, Pascal et Géraldine RIX-LIÈVRE, « Une ethnographie organisationnelle constructiviste et orientée vers les pratiques situées », *Revue internationale de psychosociologie et de gestion des comportements organisationnels*, supplément HS, 2013, p. 45-65.
- MOLLO, Valérie et Pierre FALZON, « Auto- and allo-confrontation as tools for reflective activities », *Applied Ergonomics*, vol. 35, 2004, p. 531-540.
- NIELSEN, Jakob, « The 90-9-1 rule for participation inequality in social media and online communities », *Nielsen Norman Group*, 10 septembre 2016.
- OLLION, Étienne et Julien BOELAERT, « Au-delà des *big data*. Les sciences sociales et la multiplication des données numériques », *Sociologie*, vol. 6, n° 3, 2015, p. 295-310.
- PÉLISSIER, Daniel, « Observer les observateurs, adaptation de la méthode de l'auto-confrontation simple et complexité de la navigation Internet », *Recherches qualitatives*, hors-série n° 22, 2017, p. 24-41.
- PINSKY, Leonardo et Jacques THEUREAU, « Conception des situations de travail et étude du cours d'action, un programme de recherche technologique en ergonomie », *Collection d'ergonomie et de neurophysiologie du travail*, n° 88, 1992, p. 29-82.
- PRENSKY, Marc, « Digital natives, digital immigrants, part II: Do they really think differently? », *On the Horizon*, vol. 9, n° 6, 2001, p. 1-9.
- PROST, Magalie, Béatrice CAHOUR et Françoise DÉTIENNE, « Masquage des émotions et des attentes de soutien psychologique sur des forums d'entraide professionnelle », *Activités*, vol. 10, n° 2, 2013, p. 20-38.
- RIX-LIÈVRE, Géraldine, « Différents modes de confrontation à des traces de sa propre activité, vers une confrontation à une perspective subjective située », *Revue d'anthropologie des connaissances*, vol. 4, n° 2, 2010, p. 358-379.
- ROUVROY, Antoinette, « Des données sans personne: le fétichisme de la donnée à caractère personnel à l'épreuve de l'idéologie des *Big Data* », dans *Étude annuelle*

2014 du Conseil d'État – *Le numérique et les droits fondamentaux*, La Documentation française, 2014.

SMYRNAIOS, Nikos, *Les GAFAM contre l'Internet, une économie politique du numérique*, Ina, 2017.

THEUREAU, Jacques, « Les entretiens d'autoconfrontation et de remise en situation par les traces matérielles et le programme de recherche "cours d'action" », *Revue d'anthropologie des connaissances*, vol. 4, n° 2, 2010, p. 287-322.

THEUREAU, Jacques, « L'entretien d'autoconfrontation comme composante d'un programme de recherche empirique et technologique », communication présentée aux II^{es} Journées internationales des sciences du sport, 12-15 novembre 2002.

CHAPITRE 13

La collecte manuelle des traces d'usage par la découverte progressive de mots-clics

Claudine Bonneau

Les médias sociaux renferment une quantité toujours grandissante de contenus générés par les internautes, dont l'analyse permet de découvrir et de comprendre de multiples phénomènes. Les gens qui étudient les médias sociaux procèdent généralement en délimitant le terrain à l'avance et en déterminant une communauté (des fans de hockey, par exemple), un événement (les dernières élections américaines), des comptes associés à certains types de personnes ou des mots clés, ce qui leur permet de formuler des instructions précises pouvant ensuite être opérationnalisées par des algorithmes de recherche automatisés ou de guider une collecte de données manuelle. En effet, afin de repérer les publications, les discussions, les pages et les internautes concernés par le phénomène à l'étude, des techniques de recherche simples ou de fouille de texte (*text mining*) plus sophistiquées existent. Il s'agit alors de repérer les mots fréquemment utilisés par ceux et celles qui publient du contenu lié au thème de la recherche, puis d'extraire les publications correspondantes en ajoutant d'autres paramètres de segmentation au besoin, tels que des critères sociodémographiques et temporels.

L'utilisation répandue des mots-clics (*hashtag*) sur les médias sociaux facilite la recherche thématique de publications, pour constituer un corpus à partir des publications associées à un ou à plusieurs mots-clics liés aux objectifs de recherche et aux concepts théoriques. Le terrain de recherche peut ainsi être délimité en présélectionnant des mots-clics parmi (1) ceux

promus officiellement par des associations, des marques, des personnes organisant un événement ou des influenceurs et influenceuses, (2) les plus populaires ou mis en avant par la plateforme (*trending*) ou (3) ceux qui émergent des pratiques au sein de communautés ciblées.

Cependant, certains terrains de recherche ne peuvent être circonscrits avant la collecte parce que leurs frontières sont floues, fluides et en constante reconfiguration. Une telle situation peut se présenter lorsque l'objet d'étude n'est lié ni à une communauté préexistante ni à un événement en particulier ou si les publications concernent un sujet vaste et hétérogène qui ne peut être ciblé à l'aide de mots-clés précis.

L'étude de phénomènes diffus sur les médias sociaux

Ce chapitre s'adresse aux équipes de recherche aux prises avec un tel défi, en leur proposant une démarche de collecte de données manuelle à la fois structurée et flexible. Il s'agit de l'une des stratégies possibles pour opérationnaliser l'approche de densification des données présentée par Latzko-Toth *et al.* (le présent ouvrage). Bien que le présent chapitre soit consacré à l'étape de la collecte des données et n'aborde pas les méthodes d'analyse, notons tout de même que la démarche présentée est particulièrement adaptée aux recherches qualitatives. En effet, les méthodes qualitatives sont tout indiquées pour l'exploration de phénomènes en mouvement et l'étude de pratiques en émergence qui reposent sur une compréhension exhaustive du contexte social, culturel, matériel, spatial et temporel dans lequel se déploient les usages des médias sociaux.

Ce type de recherche nécessite l'inclusion de données suffisamment détaillées et complètes pour permettre d'en dégager du sens et des significations (Onwuegbuzie et Leech, 2007). Or ces données riches et denses se prêtent plus difficilement à une analyse computationnelle. Par exemple, l'hétérogénéité et la variabilité des pratiques en ligne, notamment les cas « non standards », « extrêmes » ou « en marge » se retrouvent bien souvent ignorés, dilués ou mal interprétés par les analyses automatisées (Highfield et Leaver, 2015). De plus, on estime que 90 % de tout le contenu disponible en ligne est constitué de données non structurées (Vijayan, 2015, cité dans McKenna *et al.*, 2017). Le fait que les contenus textuels soient plus facilement collectés et traités automatiquement, comparativement aux images, a d'ailleurs pour conséquence que les recherches mobilisant les médias

sociaux sont encore peu nombreuses à s'attarder au contenu visuel comme les photos, vidéos, animations GIF, mèmes et emojis (Towner, 2018). Pourtant, les médias sociaux ont « transformé la communication humaine précisément parce que les images y occupent une place aussi importante que le texte ou la parole » (traduction libre de Miller et Sinanan, 2017, p. 1).

À l'instar de Bowker (2007), nous considérons que les contenus générés par les usagers et usagères des médias sociaux représentent en soi des traces de leurs activités en ligne (plutôt que de limiter notre définition de « traces » aux métadonnées que ces activités génèrent). L'analyse des traces en « format image » nécessite une prise en compte de leurs dimensions intertextuelles (voir Casemajor, le présent ouvrage). En effet, une image ne peut être isolée des circonstances et du contexte qui lui confèrent sa signification (Hand, 2017). Par exemple, l'image est située parmi d'autres qui l'ont précédée ou lui ont succédé au sein du profil de l'internaute qui l'a publiée, ou au sein d'une communauté qui mobilise un même mot-clic. De plus, différentes formes de communication visuelle sont employées sur les médias sociaux : photos, dessins, collages, animations GIF, mèmes, créations artistiques, citations inspirantes et poèmes présentés sous forme d'images. Chacune de ces formes vient avec ses propres conventions, accroissant ainsi son potentiel polysémique. De plus, l'image ne peut être isolée des autres éléments qui composent une publication (légende textuelle, emoji, mots-clics, commentaires), tout comme les éléments textuels ne sauraient être isolés des éléments graphiques : ils doivent donc être interprétés ensemble lors de l'analyse.

Ces considérations nous rappellent l'importance du contexte pour l'analyse des pratiques en ligne (boyd et Crawford, 2012). Alors que les techniques automatisées ont généralement pour effet d'extraire et d'isoler les traces en ligne du contexte sociomatériel de leur production, séparant ainsi la collecte et l'analyse des données, les techniques de collecte manuelle impliquent une familiarisation du chercheur ou de la chercheuse avec l'environnement étudié et leur immersion au sein de celui-ci, ce qui implique que l'interprétation ne s'effectue pas uniquement au moment de l'analyse, mais déjà au moment de la collecte des données. L'approche décrite dans ce chapitre s'inscrit donc dans une démarche ethnographique où les pratiques en ligne sont étudiées dans leur contexte « naturel » et situé, permettant ainsi de saisir les conventions culturelles ainsi que les affordances techniques qui les rendent possibles (Hine, 2015). L'ethnographie

s'appuie sur différentes méthodes de collecte qui « devraient coévoluer avec leurs objets d'étude et s'adapter continuellement aux terrains » (traduction libre de Latzko-Toth *et al.*, 2017, p. 201). En ce sens, Amit (2004, p. 17) explique justement que le fait de ne pas adapter ses pratiques sur le terrain aux nouvelles circonstances qui en émergent reviendrait à ne plus faire de l'anthropologie. Concrètement, cela signifie que le terrain ne peut être dissocié des activités qui ont permis de le délimiter, puisqu'il ne s'agit pas d'une entité préexistante qui « attend d'être découverte ». La méthode de collecte manuelle constitue donc un moyen d'opérationnaliser le processus de construction itérative d'un terrain de recherche en ligne.

À travers une problématique de recherche tirée de nos travaux, nous donnons un exemple concret de terrain ne pouvant être délimité *a priori*, pour envisager comment une stratégie de collecte manuelle peut être mise en œuvre afin de trouver ce qu'il est impossible de caractériser à l'avance de façon précise.

La mise en visibilité du travail sur Instagram

Dans le cadre d'une recherche en cours¹, nous examinons les aspects informels du travail dans le cadre des médias sociaux « grand public ». Ces derniers deviennent le lieu d'expression de réalités vécues au travail, favorisant l'émergence d'un phénomène que nous avons appelé le « travail à haute voix », où des personnes partagent de façon informelle des fragments de leur quotidien professionnel au-delà des frontières de leur organisation. D'un point de vue conceptuel, notre recherche s'inscrit au sein des études organisationnelles s'intéressant à la sociomatérialité des pratiques de travail (Orlikowski, 2007; Orlikowski et Scott, 2016) et implique donc d'examiner les possibilités qui émergent de l'intrication des acteurs et des outils numériques. Les questions qui nous animent sont les suivantes : que racontent et montrent les individus sur les médias sociaux à propos de leur travail ? Que peuvent nous apprendre ces publications sur le travail et la vie organisationnelle ?

Nous avons d'abord étudié le travail à haute voix sur la plateforme Twitter (Sergi et Bonneau, 2016). Nous avons alors remarqué que les pro-

1. L'autrice tient à remercier Viviane Sergi et Héloïse Guillaumin, respectivement coresponsable et membre de l'équipe de recherche du projet dont il est question dans ce chapitre, pour leur contribution à l'élaboration et à la documentation de l'approche présentée ici.

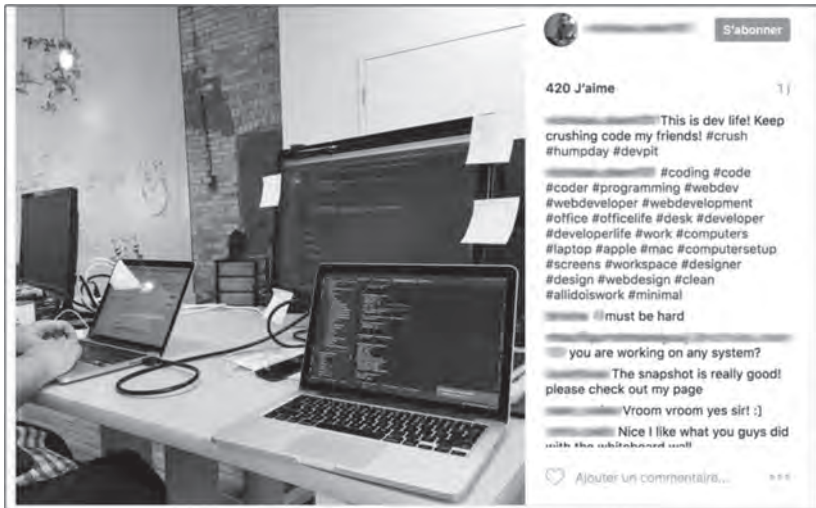
priétés performatives de ces traces d'usages composées de textes ou d'images sont étroitement liées, dans plusieurs cas, à leur dimension visuelle. La présence ubiquitaire de la photographie numérique, combinée aux fonctionnalités de partage des médias sociaux, crée un contexte propice à la diffusion instantanée d'artefacts visuels qui racontent la vie au travail. Afin d'explorer plus en détail l'aspect visuel du travail à haute voix, nous avons ensuite porté notre attention sur Instagram, une plateforme de partage de photos et de vidéos (Sergi et Bonneau, 2017). Le présent chapitre présente donc la méthode de collecte utilisée pour repérer et extraire les publications Instagram qui composent notre corpus². Même si cette approche tient compte des spécificités d'Instagram, elle peut facilement être adaptée à une collecte de données sur d'autres plateformes.

Instagram est un réseau socionumérique foncièrement visuel, puisque chaque publication doit inclure une image ou une courte vidéo. L'étiquetage des publications à l'aide de mots-clics est au cœur de l'expérience de partage d'images sur Instagram parce qu'il permet aux autres de découvrir des publications qui les intéressent, même s'ils ne sont pas abonnés au compte des personnes en question (Locatelli, 2017). Ces usagers et usagères peuvent d'ailleurs employer jusqu'à 30 mots-clics pour catégoriser et contextualiser leurs publications, en décrivant ce que représente l'image ou, à l'inverse, en communiquant des aspects qui n'y sont pas nécessairement visibles, tels que le contexte, des émotions et des idées. Par exemple, la figure 13.1 montre l'environnement de travail d'un développeur informatique. Les mots-clics qui l'accompagnent sont liés à la nature de ce type de travail en particulier (`#coding`, `#programming`), tandis que d'autres sont liés aux outils mobilisés (`#mac`, `#screens`). On y retrouve également des mots-clics relatifs à des communautés thématiques sur Instagram (`#officelife`, `#developerlife`). Alors que certains mots-clics se sont standardisés et imposés au fil des usages, d'autres sont créés spontanément pour communiquer une intention ou des conditions de travail de manière plus spécifique (`#allidoiswork`).

Une des particularités de notre recherche repose sur le caractère informel des pratiques de travail à haute voix qui émergent à l'initiative de personnes qui travaillent dans n'importe quel domaine ou dans

2. Le présent chapitre ne traite que de l'approche de collecte utilisée dans cette recherche, et non de ses résultats.

FIGURE 13.1

Exemple d'une publication Instagram liée au travail

n'importe quelle région géographique. Puisque l'un des objectifs de cette recherche est de documenter la diversité des formes que peuvent prendre ces pratiques, nous ne souhaitons pas limiter notre collecte de données à des organisations ou à des communautés professionnelles spécifiques. En effet, nous voulons être en mesure d'examiner la multiplicité des formes que le travail à haute voix peut prendre sur Instagram. Puisque nous nous intéressons aux personnes appartenant à la fois à la catégorie des travailleurs autonomes et des employés d'entreprise, dont le seul point en commun est de montrer des aspects de leur travail, quel qu'il soit, il nous est impossible de définir un point de départ *a priori*.

Ainsi, les publications Instagram qui nous intéressent ne peuvent pas être repérées en utilisant des motifs sémantiques spécifiques, puisque les sujets qui sont traités et les mots-clics utilisés varient en fonction du type d'activité professionnelle, du contexte entourant la publication et des communautés auxquelles appartiennent les personnes. Par exemple, une recherche à partir du mot-clic #accountantlife permet de repérer des publications montrant le travail de comptables, mais elle n'est pas suffisante pour documenter adéquatement la diversité des pratiques de mise en visibilité du travail sur Instagram. Cet objectif de recherche impose

donc un défi méthodologique particulier, soit le repérage de publications qui n'utilisent pas nécessairement des mots-clics identiques, puisqu'elles concernent divers types d'utilisateurs et d'utilisatrices, et de communautés. Dans la prochaine section, nous montrons comment nous avons procédé pour découvrir progressivement des personnes provenant de différents horizons techniques et professionnels, afin d'assembler un corpus de 150 publications Instagram publiques illustrant le phénomène du travail à haute voix.

La collecte manuelle de publications Instagram

La recherche et l'extraction de publications Instagram publiques peuvent se faire de plusieurs façons, par exemple, à l'aide de l'API rendue disponible par la plateforme, d'un outil externe qui se connecte à l'API d'Instagram³, en exécutant un script Python à l'aide d'un navigateur Web⁴ ou en utilisant l'interface d'utilisation de la plateforme. Lorsque ni des mots-clics ni des personnes ne peuvent être ciblés à l'avance, cette dernière option est la plus appropriée. Elle implique de trouver puis d'extraire les données manuellement en utilisant les mêmes fonctionnalités que l'utilisateur *lambda*, c'est-à-dire le moteur de recherche disponible sur la plateforme ainsi que la navigation au moyen des profils, des mots-clics, des localisations géographiques et des suggestions mises en avant par Instagram. Les désavantages occasionnés par ce processus manuel long et fastidieux sont compensés par plusieurs bénéfices. En effet, collecter les données directement à partir de l'interface d'Instagram permet de partager une expérience semblable à celle des personnes qui créent et qui consultent les publications. Les traces sont donc examinées au sein de leur environnement d'origine, plutôt que

3. Par exemple, l'application InstaR fournit une interface permettant d'accéder à l'API d'Instagram par le logiciel de programmation R, afin de télécharger les publications en fonction de mots-clics, de leur popularité, de profils ou de lieux géographiques. De tels outils externes permettent d'ajouter des fonctionnalités qui n'existent pas dans l'interface d'utilisation native, comme par exemple la possibilité de filtrer les publications en fonction de la combinaison de plusieurs critères : cran.r-project.org/web/packages/instaR/instaR.pdf

4. De tels scripts permettent de se connecter à Instagram sans passer par son API et d'en contourner certaines des limitations. Par exemple, le script Instagrab, développé par le professeur Patrik Wikström du Digital Media Research Center de la Queensland University of Technology, permet de spécifier un nombre de publications utilisant un mot-clic et de les extraire dans un fichier HTML : mementumblog.wordpress.com/2017/02/27/step-by-step-collecting-instagram-data/

d’être décontextualisées par un processus d’extraction automatisé et restituées par l’API au sein d’un fichier de données ou d’un tableur.

Notre démarche de collecte manuelle est composée des cinq étapes décrites dans les sous-sections suivantes ; une synthèse est ensuite présentée dans le tableau 13.1.

Une définition de la pratique étudiée

Puisque la collecte de données ne s’effectue pas à partir de certains mots-clics présélectionnés, on doit se doter d’autres moyens pour définir clairement ce qu’on cherche, en élaborant une définition ainsi que des critères d’inclusion et d’exclusion pouvant être facilement opérationnalisés lors de la collecte. Dans notre projet, le point de départ de la collecte s’appuie sur notre définition du travail à haute voix sur Instagram, qui s’énonce comme suit :

- Toute publication permettant de visualiser des aspects liés au travail et à l’expérience vécue du travail, dans la plupart des cas spontanément générée par la personne au moment où le travail est accompli.

Cette définition est suffisamment précise pour guider la collecte, mais volontairement inclusive pour témoigner de la diversité des pratiques étudiées et répondre aux objectifs de la recherche. Lors de notre précédente collecte sur Twitter, nous avons déjà pu désigner des mots-clics « génériques » employés par des internautes qui pratiquent le travail à haute voix sur cette plateforme (`#showyourwork`, `#shareyourwork`). Nous avons donc commencé notre exploration en vérifiant si ces mots-clics étaient aussi utilisés sur Instagram. Cependant, plutôt que d’explorer systématiquement toutes les publications trouvées à l’aide de ces mots-clics ciblés, nous avons d’abord vérifié qu’elles correspondaient bien à notre définition de départ. En effet, nous ne pouvions tenir pour acquis que les publications présentant des mots-clics liés au travail correspondaient toutes à du travail à haute voix, puisque les intentions qui motivent l’utilisation d’un mot-clic varient nécessairement d’une personne à l’autre.

Par exemple, un même mot-clic peut avoir une fonction descriptive pour une personne, alors qu’il sera plutôt utilisé à des fins ironiques, critiques ou comme pourriel par une autre. En outre, un mot-clic peut

être utilisé stratégiquement par une personne qui souhaite maximiser sa visibilité, même s'il ne correspond pas du tout au sujet de la publication. Contrairement à une collecte automatisée, le processus manuel permet donc d'évaluer la pertinence de chaque publication en regard de la définition de départ et de minimiser l'inclusion de données non valides au sein du corpus (Laestadius, 2017, p. 582).

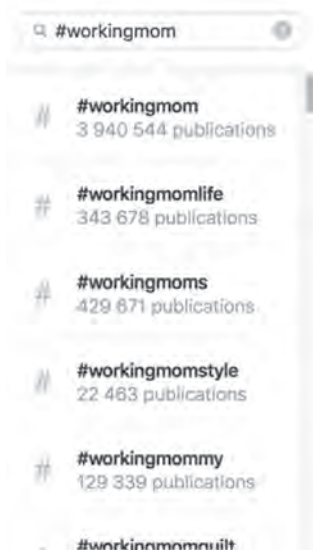
Une recherche « par empathie » pour les premiers mots-clics

La définition de départ est utilisée pour découvrir les premiers mots-clics pouvant nous conduire à des publications correspondant aux pratiques étudiées. Elle permet en quelque sorte de procéder à une recherche « par empathie », où nous tentons de nous mettre dans la peau des personnes qui pratiquent le travail à haute voix sur Instagram et en nous posant la question suivante : « Qui est susceptible de montrer son travail, et comment le décrirait-il ? » Il suffit ensuite de faire des tests dans le moteur de recherche (#hospital, #baker, #deadline, #workingmom) afin de découvrir des personnes très actives dans des domaines professionnels variés (par exemple, des travailleuses et travailleurs « manuels » comme les bouchers et les mécaniciennes, des professionnels du milieu hospitalier, des agricultrices, etc.)

Il importe également de souligner que, même s'il s'agit d'une collecte manuelle de données qualitatives, les algorithmes de la plateforme peuvent tout de même être exploités lors de la recherche de mots-clics pertinents. En effet, la fonctionnalité de recommandation du moteur de recherche peut s'avérer grandement utile pour repérer de nouveaux mots-clics. Lorsqu'on commence à taper un mot dans le champ de recherche, celui-ci nous affiche des suggestions de mots-clics connexes par ordre de popularité. La figure 13.2 montre les mots-clics affichés après avoir entré #workingmom. On peut aussi utiliser cette fonctionnalité à des fins de « bricolage », par exemple en essayant plusieurs suffixes à la suite d'un même mot-clic (#working+a, #working+b, #working+c, etc.) et en consultant les suggestions affichées.

FIGURE 13.2

Fonctionnalité de recommandation du moteur de recherche permettant de repérer des mots-clis connexes



Une stratégie « boule de neige » pour d'autres mots-clis

En examinant de près les premières publications collectées, d'autres mots-clis pertinents peuvent être repérés par effet « boule de neige »⁵. Dans une même publication, on peut retrouver des mots-clis dans plusieurs langues ou différents dérivés d'un même sujet (#work, #atwork, #working). Par exemple, nous avons trouvé des mots-clis spécifiques à Instagram (#instawork, #instajob), des mots-clis se rapportant à des domaines professionnels précis (#farmlife, #firefighter) ou mettant en avant différents contextes et des expériences de travail (#workingonsunday, #toomuch-

5. L'effet « boule de neige » fait ici référence à une méthode d'échantillonnage non probabiliste permettant de repérer de nouveaux sujets à partir des informations fournies par d'autres sujets déjà identifiés (Noy, 2008). La métaphore de la boule de neige évoque la dimension cumulative du processus qui permet de constituer l'échantillon progressivement, c'est-à-dire que les premiers sujets sélectionnés en suggèrent d'autres qui, à leur tour, en suggéreront de nouveaux. En ce qui concerne plus particulièrement la collecte de données à l'aide des médias sociaux, il s'agit de mobiliser les relations, les liens et les interactions rendus visibles par les plateformes afin de trouver des sujets à inclure dans l'étude (Baltar et Brunet, 2012).

work). Le fait d'utiliser les publications déjà répertoriées pour la suite du repérage permet de ne pas négliger des mots-clics plus marginaux, controversés (#fuckmyjob) ou renfermant des fautes d'orthographe. Ceux-ci peuvent nous mener à d'autres publications pertinentes qui n'auraient pu être repérées en se concentrant uniquement sur les mots-clics les plus populaires.

L'immersion continue dans les pratiques de travail à haute voix permet de faire des déductions sémiotiques à partir des publications déjà collectées. En effet, l'interprétation du contenu visuel de la photo peut faire penser à d'autres mots-clics à essayer dans le moteur de recherche. Par exemple, une photo représentant l'environnement de travail d'un technicien biomédical nous incite à mettre à l'essai des mots-clics liés aux pratiques et aux outils mobilisés par des professionnels de laboratoire. Aussi, la découverte de mots-clics reflétant le genre (#tradergirl) nous amène à essayer certains mots-clics en version féminine.

La collecte des publications sans mots-clics

Certaines personnes n'utilisent aucun mot-clic, soit parce qu'elles n'y voient pas d'intérêt, soit pour limiter stratégiquement la portée de leurs publications. Dans notre cas, certaines publient des photos à l'insu de leur employeur ou rendent visibles certains aspects délicats, voire confidentiels de leur travail à partir d'un pseudonyme et réduisent la portée de leurs publications – même si elles sont publiques – en n'utilisant pas de mot-clic. Afin d'élargir la collecte aux publications sans mots-clics, il s'agit, par exemple, d'utiliser les publications déjà intégrées au corpus afin de repérer d'autres publications (1) qui sont géolocalisées au même endroit, (2) qui proviennent d'abonnements à un compte déjà intégré au corpus ou (3) des profils suivis, ou encore (4) de personnes les ayant commentées. Cette autre forme de navigation «boule de neige» permet, dans certains cas, de percevoir les contours de certaines communautés thématiques au-delà des mots-clics.

D'autres méthodes de collecte de données

Nous insistons également sur l'utilité, voire la nécessité, de compléter la collecte de traces en ligne – qu'elle soit effectuée manuellement ou automatiquement – par d'autres méthodes permettant d'aborder les pratiques

en ligne avec les personnes étudiées afin de valider les interprétations et de documenter les significations qu'elles attribuent à leurs usages. Dans notre cas, l'utilisation de méthodes d'enquête complémentaires – telles que l'entrevue sur traces décrite par Latzko-Toth *et al.* (le présent ouvrage) – s'avère nécessaire pour comprendre le contexte plus large dans lequel s'insère le travail à haute voix, notamment pour savoir s'il s'effectue à l'insu ou non de l'employeur, si les interactions qui en ont découlé ont permis de tisser de nouveaux liens ou d'établir des collaborations, etc. Ces stratégies complémentaires peuvent également servir à accéder à ce qui échappe à l'analyse des traces, notamment les activités qui ne laissent aucune trace en ligne, comme la décision de ne pas publier une photo, par exemple. Comme la plupart des phénomènes humains pouvant être étudiés à l'aide des traces d'usage sur les médias sociaux, le travail à haute voix ne saurait être réduit à ce qui se déroule en ligne.

Une synthèse

Le tableau 13.1 présente une synthèse des différentes étapes de la méthode de collecte manuelle proposée ainsi que des trucs et des conseils destinés à guider leur opérationnalisation. Nous soulignons l'importance, tout au long du processus de collecte, de documenter de façon claire et détaillée les règles élaborées ainsi que les chemins empruntés, afin de pouvoir les justifier adéquatement lors de la publication des résultats et de guider l'équipe de recherche tout au long de l'étude. Nous recommandons également l'utilisation d'un outil de capture d'écran pour sauvegarder la publication dans son ensemble (et non seulement le fichier de l'image) et son insertion dans un journal de bord renfermant l'URL, la date de la publication et la façon dont elle a été repérée.

Finalement, la collecte manuelle constitue non seulement une occasion de se familiariser avec son terrain d'étude, mais aussi d'affiner progressivement sa compréhension du phénomène à l'étude et de voir émerger des modèles, et ce, avant même d'entamer formellement l'analyse. Dans notre cas, ce processus manuel alliant flexibilité, sérendipité et bricolage nous a permis de découvrir des pratiques inattendues ou contre-intuitives qui n'auraient pas pu être repérables autrement, tout simplement parce que nous n'aurions pas pu formuler des critères de recherche appropriés avant de savoir qu'elles existaient. Bien que la collecte manuelle ne puisse être utilisée

TABLEAU 13.1

Vue d'ensemble de la méthode de collecte manuelle proposée

Étapes	Trucs et conseils
1. Définir la pratique étudiée et préciser les critères d'inclusion et d'exclusion.	Utiliser la question et les objectifs du projet pour déterminer « ce que l'on cherche » et confirmer la validité des données recueillies.
2. Effectuer une recherche « par empathie » afin de trouver les premiers mots-clics.	Entrer des termes suggérés par la question de la recherche afin de repérer des publications renfermant certains mots-clics. Utiliser la fonctionnalité de recommandation du moteur de recherche pour repérer des mots-clics associés.
3. Appliquer une stratégie « boule de neige » pour repérer d'autres mots-clics.	Utiliser les publications déjà répertoriées pour la suite du repérage, en examinant les autres mots-clics qu'elles renferment et en effectuant des déductions sémiotiques en fonction de l'interprétation du contenu visuel.
4. Élargir la collecte aux publications sans mots-clics.	Utiliser les publications déjà répertoriées afin de repérer d'autres publications 1) géolocalisées au même endroit, 2) provenant des abonnements associés au profil déjà intégré au corpus, 3) des comptes qu'il suit et 4) des personnes les ayant commentées.
5. Mobiliser d'autres méthodes pour compléter la collecte de données.	Par exemple, l'entrevue sur traces (voir Latzko-Toth <i>et al.</i> , le présent ouvrage) permet d'aborder les pratiques en ligne avec les personnes participantes elles-mêmes et de densifier les données collectées.

pour constituer de larges échantillons, elle peut être employée stratégiquement pour amasser des informations précieuses sur les contextes d'usage et les pratiques entourant la production de ces publications, permettant ainsi de densifier les traces recueillies. De plus, elle permet de développer une conscience et une connaissance fine des affordances et des normes de la plateforme en question. On peut alors plus aisément se tenir au courant des changements techniques, comprendre les nouvelles fonctionnalités qui surgissent au fil de l'évolution de la plateforme et déceler les implications de ces évolutions pour les résultats de recherche. En ce sens, les méthodes de collecte manuelle permettent de développer une forme de dextérité dans la manipulation de traces multimodales en contexte. Ces habiletés sont plus que jamais nécessaires pour l'étude des modes de communication « post-texte » qui caractérisent les usages des dispositifs mobiles et des médias sociaux visuels tels qu'Instagram, Snapchat et Tik Tok.

La relation d'intimité et de familiarité entre le chercheur ou la chercheuse et son sujet constitue le moyen fondamental de recherche qui distingue les études ethnographiques d'autres types de méthodes (Amit, 2004, p. 2). Or, dans les environnements numériques en constante reconfiguration, le sentiment de familiarité développé par rapport à son terrain ne peut être que temporaire. L'agilité et la résilience – des compétences hautement demandées en recherche et dans d'autres domaines – constituent aussi des atouts permettant de composer avec des plateformes et des pratiques caractérisées par l'impermanence. Les méthodes de recherche qualitatives sont donc appelées à évoluer dans le sens d'une meilleure malléabilité face aux environnements changeants.

Références

- AMIT, Vered (dir.), *Constructing the Field: Ethnographic Fieldwork in the Contemporary World*, Routledge, 2004.
- BALTAR, Fabiola et Ignasi BRUNET, « Social research 2.0: Virtual snowball sampling method using Facebook », *Internet Research: Electronic Networking Applications and Policy*, vol. 22, n° 1, 2012, p. 57-74.
- BONNEAU, Claudine, « Travailler à haute voix sur Twitter: Quand la collaboration informelle emprunte un réseau public », *tic&société*, vol. 6, n° 2, 2013.
- BOWKER, Geoffrey C., « The past and the Internet », dans Joe KARAGANIS (dir.), *Structures of Participation in Digital Culture*, Social Science Research Council, 2007, p. 20-36.
- BOYD, danah et Kate CRAWFORD, « Critical questions for big data: Provocations for a cultural, technological and scholarly phenomenon », *Information, Communication & Society*, vol. 15, n° 5, 2012, p. 662-679.
- HAND, Martin, « Visuality in social media: Researching images, circulations and practices », dans Anabel QUAN-HAASE et Luke SLOAN (dir.), *The SAGE Handbook of Social Media Research Methods*, SAGE Publications, 2017, p. 215-231.
- HIGHFIELD, Tim et Tama LEAVER, « A methodology for mapping Instagram hashtags », *First Monday*, vol. 20, n° 1, 2015.
- HINE, Christine, *Ethnography for the Internet: Embedded, Embodied and Everyday*, Bloomsbury Publishing, 2015.
- LAESTADIUS, Linnea, « Instagram », dans Anabel QUAN-HAASE et Luke SLOAN (dir.), *The SAGE Handbook of Social Media Research Methods*, SAGE Publications, 2017, p. 573-592.
- LATZKO-TOTH, Guillaume, Claudine BONNEAU et Mélanie MILLETTE, « Small data, thick data: Thickening strategies for trace-based social media research », dans Anabel QUAN-HAASE et Luke SLOAN (dir.), *The SAGE Handbook of Social Media Research Methods*, SAGE Publications, 2017, p. 199-214.

- LOCATELLI, Elisabetta, « Images of breastfeeding on Instagram : Self-representation », *Social Media + Society*, vol. 3, n° 2, 2017, p. 1-14.
- MCKENNA, Brad, Michael D. MYERS et Michael NEWMAN, « Social media in qualitative research: Challenges and recommendations », *Information and Organization*, vol. 27, n° 2, 2017, p. 87-99.
- MILLER, Daniel et Jolynna SINANAN, *Visualising Facebook*, UCL Press, Londres, 2017.
- NOY, Chaim, « Sampling knowledge: The hermeneutics of snowball sampling in qualitative research », *International Journal of Social Research Methodology*, vol. 11, n° 4, 2008, p. 327-344.
- ONWUEGBUZIE, Anthony J. et Nancy L. LEECH, « Validity and qualitative research: An oxymoron? », *Quality & Quantity*, vol. 41, n° 2, 2007, p. 233-249.
- ORLIKOWSKI, Wanda J., « Sociomaterial practices: Exploring technology at work », *Organization Studies*, vol. 28, n° 9, 2007, p. 1435-1448.
- ORLIKOWSKI, Wanda J. et Susan V. SCOTT, « Digital work: A research agenda », dans Barbara CZARNIAWSKA (dir.), *A Research Agenda for Management and Organization Studies*, Edward Elgar Publishing, Northampton, 2016, p. 88-96.
- SERGI, Viviane et Claudine BONNEAU, « As I see life at work: Sharing work experiences on social media », dans 33rd EGOS Colloquium, *Sub-theme 61: Viewing the Unseen Organization in Practice*, Copenhagen, 2017.
- SERGI, Viviane et Claudine BONNEAU, « Making mundane work visible on social media: A CCO investigation of working out loud on Twitter », *Communication Research and Practice*, vol. 2, n° 3, 2016, p. 378-406.
- TOWNER, Terri L., « Image is the new text: The rise of digital visuals in communication », *The Electronic Journal of Communication*, appel de textes, 2018.
- VIJAYAN, Jaikumar, « Solving the unstructured data challenge », *CIO Magazine*, 25 juin 2015.

QUATRIÈME PARTIE

**Les méthodes d'analyse d'images
et les dispositifs vidéo**

CHAPITRE 14

L'analyse visuelle du Web : une approche forensique des métadonnées d'images

Nathalie Casemajor

En 2012, le magazine en ligne *Vice* a publié un article sur John McAfee, le créateur millionnaire du logiciel antivirus homonyme, en cavale et recherché par la justice américaine. Le journaliste s'est déplacé en Amérique centrale pour faire une entrevue avec lui dans un lieu tenu secret. Mais dans la photographie publiée avec l'article en ligne se trouvait une balise de géolocalisation, pointant vers un restaurant au Guatemala. Les métadonnées intégrées au fichier image ont donc involontairement révélé l'emplacement du fugitif. Les métadonnées d'images sont un ensemble d'informations qui documentent la production, le traitement et la circulation des images dans les systèmes informatisés. Si elles sont des opérateurs de fonctionnement des systèmes informatiques (elles permettent de localiser et de traiter les images), ces métadonnées archivent aussi des traces d'activité enregistrées automatiquement, parfois à l'insu des personnes concernées. Elles donnent ainsi des indices sur le cycle de vie d'un artefact visuel, éclairant les usages, les pratiques sociales et les infrastructures techniques qui les entourent.

Ce chapitre¹ traite des métadonnées d'images et de leur exploitation dans le cadre d'une ethnographie visuelle du Web. Il examine les possibilités offertes par ces métadonnées du point de vue méthodologique, tout en évaluant leurs limites et les enjeux éthiques qu'elles soulèvent. L'objet de cette réflexion concernera plus particulièrement les métadonnées « intégrées » aux fichiers image (*embedded metadata*), c'est-à-dire les informations qui sont stockées à l'intérieur même du fichier, et non dans une base de données associée. Peu de travaux se sont intéressés aux modalités de leur collecte et de leur analyse. Le statut ambigu de ces données dans le régime de visibilité numérique soulève des considérations éthiques liées à leur exploitation en recherche. Quels sont le potentiel et les limites des métadonnées intégrées aux images pour étudier l'activité sociale en ligne, les modèles économiques des plateformes et les rapports de pouvoir sur le Web? Ces métadonnées sont comme des microarchives mobiles qui posent plusieurs défis méthodologiques. En abordant les enjeux éthiques de ce matériau d'enquête sous l'angle de la responsabilité épistémique des chercheurs et chercheuses qui les exploitent, nous proposons la notion de « visibilité contextuelle » des métadonnées sur les interfaces, de façon à mettre en lumière la responsabilité partagée des responsables des recherches et des plateformes dans la gestion des données sensibles.

Les enjeux méthodologiques des images et métadonnées intégrées

Les métadonnées intégrées sont un ensemble d'informations descriptives, techniques et administratives conservées dans les fichiers images (mais aussi dans les fichiers sonores, vidéo et textuels). Stockées à l'intérieur des segments de code qui composent l'architecture du fichier lui-même, elles décrivent la structure et les attributs des ressources électroniques.

Des microarchives mobiles

Ce qui rend les métadonnées intégrées particulièrement intéressantes du point de vue de la recherche, c'est qu'elles opèrent comme une micro-

1. Merci à Karine Michaud-Tessier pour sa contribution à la traduction, à Natasha Chaykowski et Rosa Iris Rodriguez Rovira pour leur contribution à la collecte de données de recherche, et aux membres du groupe de lecture Bits, Bots et Bytes coordonné par E. Gabriella Coleman à l'Université McGill pour leurs commentaires éclairants. Cette recherche a bénéficié d'un financement de l'Université du Québec en Outaouais.

archive mobile. Voyageant dans le fichier au fur et à mesure de son déplacement d'une plateforme à l'autre, elles documentent le cycle de vie des images. La figure 14.1 montre une série de métadonnées extraites d'un fichier JPEG. On y trouve par exemple la marque de l'appareil photo utilisé, son numéro de série, la date et l'heure de création de la photographie, le nom de la personne qui l'a produite, les coordonnées de localisation géographique où la photo a été prise, la date à laquelle le fichier a été modifié et le nom de l'application utilisée. L'ensemble de ces traces de prise de vue et d'usage du fichier composent la biographie de l'image, c'est-à-dire qu'elles racontent l'histoire de ce fichier du point de vue de son inscription dans différents systèmes d'enregistrement et de gestion de l'information.

FIGURE 14.1

Métadonnées extraites d'une photographie publiée sur Flickr au sujet de la grève étudiante de 2012 au Québec. Stéphane Gauthier, «Red squares are everywhere. Cathédrale Marie-Reine-du-Monde», 2012.

```

EXIF | Make - Canon | X-Resolution - 72 dpi | Y-Resolution - 72 dpi | Software - Adobe Photoshop Lightroom 4.1 (Windows)
| Date and Time (Modified) - 2012:06:03 22:19:56 | Artist - Stéphane Gauthier | Copyright - ©Stéphane Gauthier | ISO Speed
- 200 | Exif Version - 0230 | Date and Time (Original) - 2012:05:31 17:48:31 | Date and Time (Digitized) - 2012:05:31
17:48:31 | Exposure Bias - 0 EV | Metering Mode - Multi-segment | Sub Sec Time Original - 13 | Sub Sec Time Digitized - 13 |
Focal Plane X-Resolution - 3840.211789 | Focal Plane Y-Resolution - 3908.141962 | Focal Plane Resolution Unit - inches |
Custom Rendered - Normal | Exposure Mode - Auto | White Balance - Auto | Scene Capture Type - Standard | Owner Name -
Stéphane Gauthier | GPS Latitude Ref - North | GPS Latitude - 45 deg 29.27; 58.53 | GPS Longitude Ref - West | GPS
Longitude - 73 deg 34.27; 9.03 | Compression - JPEG (old-style) | Thumbnail Offset - 888 | Thumbnail Length - 22834
Photoshop | Displayed Units X - inches | Displayed Units Y - inches | Copyright Flag - True | Photoshop Thumbnail - (Binary
data 22834 bytes, use -b option to extract) | IPTCDigest - fcd534204014f9666adb84cf43627733 |
IPTC | Coded Character Set - UTF8 | Application Record Version - 4 | Special Instructions - montreal-ville-marie | Date Created
- 2012:05:31 | Time Created - 17:48:31 | Digital Creation Date - 2012:05:31 | Digital Creation Time - 17:48:31 | By-line -
Stéphane Gauthier | By-line Title - Photographe | City - Montréal | Sub-location - Ville-Marie | Province-State - Québec
| Country - Primary Location Name - Canada | Copyright Notice - ©Stéphane Gauthier
XMP | XMPToolkit - Adobe XMP Core 5.3-c007.1.136881, 2010:06:10-18:11:35 | Format - image/jpeg | Creator - Stéphane
Gauthier | Rights - ©Stéphane Gauthier | Image Number - 0 | Flash Compensation - 0 | Owner Name - Stéphane Gauthier |
Firmware - 2.1.1 | Creator Tool - Adobe Photoshop Lightroom 4.1 (Windows) | Metadata Date - 2012:06:03 22:19:56
4:00 | Authors Position - Photographe | Instructions - montreal-ville-marie | City - Montréal | State - Québec | Country -
Canada | Document ID - xmp.did:46AF8EB4EBADE119B3AEA67F1A259CC | Original Document ID -
FA387927372AA605B1A7B0850E933D231 | Instance ID - xmp:did:46AF8EB4EBADE119B3AEA67F1A259CC | History
Action - derived | History Parameters - converted from images-canon-cr2 to image;dag, saved to new location | History
Instance ID - xmp:did:017298A29FABE111ABAFBBB3EB6DDEF3 | History When - 2012:06:01 00:09:51-04:00 | History
Software Agent - Adobe Photoshop Lightroom 4.0 (Windows) | History Changed - | Derived From Instance ID -
xmp:did:017298A29FABE111ABAFBBB3EB6DDEF3 | Derived From Document ID
xmp:did:017298A29FABE111ABAFBBB3EB6DDEF3 | Derived From Original Document ID -
FA387927372AA605B1A7B0850E933D231 | Marked - True | Usage Terms - Tous droits réservés. All rights reserved | Location
- Ville-Marie | Creator Region - Québec | Creator Country - Canada | Creator Work Email - contact@photographegauthier.com |
Creator City - Montréal | Creator Work URL - www.photographegauthier.com |
ICC | Viewing Cond Illuminant - 19.4445 20.3718 16.8089 | Viewing Cond Surround - 5.92859 4.07439 3.36179 | Viewing
Conditions Illuminant Type - D50 | Measurement Observer - CIE 1931 | Measurement Backing - 0 0 0 | Measurement Geometry -
Unknown (0) | Measurement Flare - 0.999% | Measurement Illuminant - D65 |
APP14 | DCTEncode Version - 100 | APP14 Flags0 - Encoded with Blend=1 downsampling | APP14 Flags1 - (none) | Color
Transform - YCbCr
MakerNotes | Camera ID - 72157607348761901 | Camera Type - Digital SLR

```

Dans l'environnement numérique, la fonction première des métadonnées est de permettre la gestion des ressources électroniques en facilitant les opérations de stockage, de classification, de traitement, d'interopérabilité, de récupération et de préservation (Dempsey et Heery, 1998). Plus généralement, les métadonnées de documents renseignent six principales catégories d'information : le type de média, la description du contenu, la classification du contenu, la composition du document, son historique et son emplacement (Bohm et Rakow, 1994). Complémentaires à l'analyse, les métadonnées transactionnelles incluent quant à elles les traces de connexion à une ressource en ligne, notamment l'adresse IP de l'internaute qui a consulté une image sur le Web, ou le nombre de fois qu'elle a été téléchargée.

L'intégration des métadonnées dans le fichier image dépend des formats de compression (JPEG, TIFF, PNG, GIF ou autre), qui ont chacun leur propre manière de structurer les schémas de métadonnées. L'information stockée sous la forme de métadonnées est principalement textuelle, mais on peut aussi y trouver des images sous la forme de vignettes qui peuvent atteindre jusqu'à 640 x 480 pixels. Ces vignettes, créées par les appareils photographiques ou les téléphones cellulaires, permettent la prévisualisation des images sur les interfaces de gestion de fichiers.

Outre la diversité des formats de compression, il existe plusieurs standards de métadonnées établissant la structure interne de l'information. Ces standards ont été développés pour adapter les modèles économiques aux mutations numériques dans les secteurs de la photographie commerciale, des agences de presse, de l'édition et de l'industrie des logiciels. Plus largement, ils témoignent de l'histoire des télécommunications et des échanges internationaux d'information depuis les années 1970, après l'émergence des téléimprimeurs et des ordinateurs personnels. Parmi les principaux acteurs qui ont influencé la mise en place de ces standards figurent l'IPTC (International Press and Telecommunication Council, fondé en 1965 pour développer les normes de l'industrie) et la société de logiciels Adobe.

Le potentiel pour l'ethnographie du Web

Ce chapitre envisage le potentiel des métadonnées d'images pour l'ethnographie du Web en s'inspirant de la démarche forensique. Le néologisme « forensique » est tiré de l'anglais *forensic* (dérivé du latin *forum*,

relatif à la place publique), utilisé en criminologie pour désigner les moyens scientifiques et les techniques employés pour révéler des indices invisibles à l'œil nu. Ce terme a été introduit dans le domaine de l'archéologie des médias numériques depuis les années 2000 pour caractériser l'extraction et l'investigation des traces numériques contenues dans des fichiers, des systèmes informatiques ou des ensembles de données (Kirschenbaum, 2008). Les travaux dans ce domaine ont contribué à mettre l'accent sur la matérialité des dispositifs numériques, à l'encontre des discours sur la dématérialisation (Casemajor, 2015).

L'approche forensique des métadonnées intégrées aux images ouvre donc une nouvelle voie pour l'ethnographie visuelle du Web en proposant d'étudier la structure interne des fichiers images. Il ne s'agit pas ici d'une ethnographie visuelle au sens de l'étude du contenu graphique des images et de l'analyse sémiotique, mais d'un matériau qui permet de saisir la dimension sociotechnique du régime visuel numérique. Les travaux qui s'inscrivent dans ce domaine contribuent à renouveler la réflexion sur le statut de l'image (photographique en particulier) entre surface visuelle et enregistrement de traces d'activité : non seulement comme captation (empreinte lumineuse) d'une scène qui a eu lieu, mais aussi comme support d'accumulation d'indices d'activité qui témoignent du cycle de vie de l'artefact numérique.

Ainsi les métadonnées d'images peuvent-elles entrer dans la composition des stratégies méthodologiques pour l'ethnographie visuelle du Web : elles permettent de comprendre l'articulation entre infrastructure technique et dimensions socioculturelles des pratiques de l'image en contexte numérique. En retraçant la trajectoire biographique d'un artefact numérique (les événements et processus qui ont marqué son cycle de vie), elles renseignent sur ses conditions de production et de circulation : qui a créé une image, quand, où, dans quel contexte, comment elle a été éditée et classifiée, où elle a été stockée, et quels sont les droits d'auteur associés. Si elles permettent d'observer les chaînes d'action entre la production et l'usage des images, les métadonnées intégrées documentent également les interactions entre agents non humains au sein des systèmes informatisés (traces d'activités automatisées).

Les domaines d'application de cette approche sont variés. Ils incluent d'une part l'étude de la provenance et de la chaîne de propriété d'un artefact dans le domaine des sciences de l'information. D'autre part, cette

approche peut être mise à profit pour l'étude des transformations des pratiques de l'image en contexte de production artistique, journalistique ou amateur dans le cadre du partage quotidien d'images dans les médias sociaux. Alors que de nombreux travaux portent plutôt sur des corpus de petite ou de moyenne taille, Hochman et Manovich (2013) expérimentent des méthodes d'analyse de corpus de grande taille. Bien que peu exploré pour l'instant du point de vue de l'archéologie des images, le domaine émergent des études des archives du Web offre également un nouvel horizon pour étudier des mégacorpus d'images dans une perspective diachronique, de façon à saisir l'évolution des pratiques de publication d'images et l'évolution de leurs conditions techniques de production et de diffusion sur le Web (Brügger, 2018).

Enfin, une des applications les plus prometteuses concerne l'étude de l'infrastructure technique des régimes de visibilité numérique : comprendre comment les fichiers image sont traités, classés, stockés et distribués par les plateformes offre un point d'entrée original pour étudier l'économie politique des infrastructures numériques. Cette approche peut nourrir les travaux qui s'intéressent aux rapports de production, de propriété et de pouvoir dans l'environnement numérique. Ces derniers concernent des enjeux tels que les luttes économiques et symboliques au sein de l'industrie de l'image et de l'information pour imposer un standard de métadonnées, les conflits autour de la suppression des métadonnées d'origine par les réseaux socionumériques (au profit de leurs propres métadonnées de gestion des images), le degré de portabilité des métadonnées d'une plateforme à une autre, la vie privée et le partage de données sensibles contenues dans les métadonnées. Cette approche méthodologique peut aussi être utile aux personnes travaillant dans la production et diffusion d'images pour étudier les usages, les modes d'appropriation et les significations créées au regard des métadonnées associées à leurs images. En somme, l'évolution des systèmes de métadonnées d'images témoigne plus largement des grandes transformations des modes de production et de circulation de l'image dans l'environnement numérique.

Les défis méthodologiques

Le tableau 14.1 décrit les quatre principaux standards de métadonnées d'images et de leurs caractéristiques. Il s'agit des standards Exif, XMP,

IPTC et ICC. Certaines de ces métadonnées sont facilement accessibles, car elles sont affichées par l'interface des systèmes d'exploitation des ordinateurs, dans des applications comme Adobe Bridge et Lightroom, ou sur des plateformes de partage de photographies en ligne telles que Flickr. Mais beaucoup ne sont pas rendues visibles sur ces interfaces. Pour accéder à l'ensemble des métadonnées contenues dans un fichier image, il faut donc utiliser un outil d'extraction. L'outil le plus couramment utilisé est le programme ExifTool, créé par Phil Harvey. Ce programme gratuit, à code source ouvert, peut être exécuté à l'aide du terminal de lignes de commande d'un ordinateur pour récupérer, afficher et modifier divers types de métadonnées dans des fichiers image, textuels ou vidéo. Plusieurs services en ligne, fondés sur ExifTool, fournissent également des outils simplifiés pour afficher les métadonnées intégrées².

Chaque catégorie de métadonnées présente des limites du point de vue de ses possibilités de collecte, de traitement et d'analyse. Une connaissance approfondie du contexte de création des images ainsi qu'une analyse complémentaire du paratexte (titres, légendes, hyperliens) sur les pages Web où elles ont été publiées peuvent contribuer à réduire le risque d'interprétations erronées. Mais une analyse approfondie n'est souvent possible que dans le cas des petits et moyens corpus. L'analyse des corpus massifs pose donc des défis supplémentaires, illustrés par les limitations suivantes. Premièrement, la date et l'heure contenues dans les métadonnées Exif peuvent être erronées si les réglages initiaux des appareils photographiques sont inexacts, ou quand l'appareil voyage sans prendre en compte la coordination entre le fuseau horaire local et le temps universel coordonné (UTC). Deuxièmement, le fait de copier ou de déplacer un fichier à l'intérieur du système d'un ordinateur peut modifier certaines métadonnées, comme la date de création du fichier. Par ailleurs, la création de copies d'écran détache complètement l'image de son bagage de métadonnées embarquées, dans la mesure où cette opération crée un nouvel objet qui capture la grille de pixels sans reproduire l'architecture de codage interne du fichier. Et troisièmement, certaines plateformes telles que Facebook, Twitter et Instagram suppriment une partie des métadonnées intégrées. Si le respect de la vie privée est une des raisons invoquées pour effacer ces

2. Voir le site *Jeffrey's Exif Viewer*: exif.regex.info/. Voir aussi le site *FotoForensics.com*, créé par Neal Krawetz de Hacker Factor, un analyste de la sécurité informatique spécialisé dans le profilage en ligne et les technologies anti-anonymat.

TABLEAU 14.1

Principaux standards de métadonnées d'image et exemples de champs

Standards de métadonnées	Type d'information enregistrée
<p>EXIF (Exchangeable Image File Format) <i>Information technique sur les conditions de création des images</i></p> <p>Générée automatiquement par les appareils photo</p> <p>Compatible avec les formats JPEG, TIFF</p>	<ul style="list-style-type: none"> • Date et heure de création • Géolocalisation (coordonnées GPS: latitude, longitude, altitude) • Nom du fabricant de l'appareil • Numéro d'identification unique de l'appareil • Réglages de l'appareil lors de la prise de vue : conditions de lumière (soleil / nuages); <i>flash on / off</i> • Distance entre l'appareil et le sujet photographié • Dimensions de l'image originale (si l'image a été recadrée ou réduite) • Vignette
<p>XMP (Extensible Metadata Platform) <i>Information technique + légale et administrative</i></p> <p>Générée par les applications (par ex., Photoshop)</p> <p>Compatible avec les formats JPEG, TIFF, GIF, PNG, PDF, MP3...</p>	<ul style="list-style-type: none"> • Histoire des modifications du fichier (date et heure où le fichier a été sauvegardé ou converti; un long historique des modifications indique qu'une image a été manipulée) • Dépistage de l'outil (nom et version du logiciel utilisé pour éditer l'image) • Numéros d'images sources (quand plusieurs images ont été combinées) • Dimensions de l'image source (si l'image a été recadrée ou redimensionnée) • Reproduction des champs Exif • Reproduction des champs IPTC
<p>IPTC (International Press Telecommunications Council) <i>Information légale et administrative</i></p> <p>Générée manuellement (fins commerciales et protection des droits d'auteur)</p> <p>Compatible avec JPEG, TIFF, PNG</p>	<ul style="list-style-type: none"> • Crédits • Titre • Légende • Mots clés • Source • Licence
<p>Profils ICC (International Color Consortium) <i>Information technique</i></p> <p>Générée par les applications (définit les paramètres de restitution des couleurs de l'image à l'écran)</p> <p>Compatible avec tous les fichiers images</p>	<p>Description du profil de couleur:</p> <ul style="list-style-type: none"> • nom du profil • type d'appareil utilisé • nom du fabricant <p>Par exemple, Facebook peut intégrer son propre profil de couleur dans les images qui transitent par sa plateforme.</p>

données, un autre avantage pour les plateformes consiste à alléger les fichiers et à faciliter le stockage et l'affichage des images en ligne, tout en générant un système de métadonnées propres à la plateforme.

Ainsi, lors de leur trajectoire de circulation sur le Web, les images sont soumises à un double processus d'accumulation et de dégradation des métadonnées intégrées : d'un côté, les interactions avec les images ajoutent des traces dans l'archive mobile du fichier, tandis que de l'autre côté, le transit par les réseaux sociaux tend à réduire la quantité et la qualité de ces traces. Les principales limites de l'analyse des métadonnées intégrées aux images tiennent donc au caractère fragmenté de ces corpus et à leur intégrité souvent lacunaire en raison de défauts dans les paramètres des appareils photo, d'opérations de manipulation des fichiers et du type de paramètres de gestion des métadonnées dans les plateformes Web.

Les enjeux éthiques de l'analyse des traces forensiques

L'essor fulgurant des pratiques de publication de photographies sur le Web, en particulier dans les médias socionumériques, pose de nouveaux défis en termes de protection de la vie privée et de consentement à l'exploitation des traces d'activité en ligne. Les modèles d'affaires des industries de l'information, qui commercialisent ces données en passant des contrats d'utilisation souvent opaques, ont fait l'objet de nombreuses critiques. Dans ce contexte, les méthodes de recherche qui exploitent les traces d'usage des contenus numériques gagnent à clarifier les principes éthiques qui guident la collecte, le traitement et la conservation de ces données. Les chercheurs et les chercheuses sont en effet prises dans des relations de pouvoir et de savoir par rapport aux objets/sujets de leur étude. Plus largement, le passage par une démarche réflexive sur le processus d'enquête peut être une source d'enrichissement du cadre méthodologique et conceptuel, ouvrant de nouvelles pistes d'analyse des dynamiques de visibilité et d'invisibilité des données.

L'évaluation des risques

De nombreux internautes ne sont pas conscients du type d'information qui peut être dissimulé dans les métadonnées d'un fichier image. Même les plus experts peuvent commettre des erreurs : en 2012, un membre du collectif de hackers CabinCr3w a voulu narguer les autorités en publiant

en ligne une photographie... qui contenait des balises de géolocalisation, ce qui a mené à son arrestation (Goodin, 2012). Si toute recherche ne présente pas nécessairement de risques, et s'il convient de relativiser ceux-ci dans le cas de corpus aussi fragmentaires que ceux des métadonnées intégrées aux images, certains terrains, portant notamment sur la production visuelle des mouvements sociaux, des jeunes et des groupes vulnérables ou marginalisés, exigent que l'on porte une attention particulière aux conséquences potentielles de la recherche.

Dans le cas des métadonnées intégrées aux images, les principaux champs pouvant contenir des informations personnelles ou sensibles sont les « balises de géolocalisation », comme il est illustré par les exemples *supra*, bien que leur présence dans les photographies en ligne soit plutôt rare³; le « numéro d'identification *unique* » d'un appareil photo, qui peut permettre de relier une série d'images à une même source; l'« historique des modifications » d'un fichier, qui peut indiquer quand une image a été manipulée; et les « vignettes », qui peuvent révéler la version originale d'une image avant qu'elle n'ait été manipulée. Murdoch et Dornseif (2004) ont par exemple montré qu'une vignette peut révéler l'identité d'une personne ayant recadré son portrait avant de le publier en ligne (figure 14.2). Dans l'échantillon d'images qu'ils ont analysé, environ 9 % des vignettes intégrées présentaient des différences avec l'image affichée à l'écran.

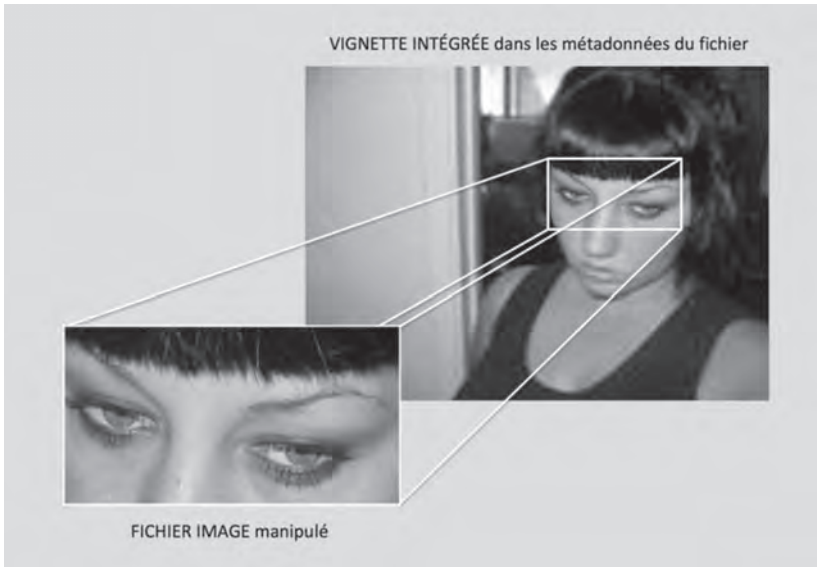
Plus généralement, les chances de trouver des métadonnées compromettantes dans une photographie dépendent du format du fichier et de la plateforme où elle a été publiée. JPEG, le standard d'images le plus commun sur le Web, peut contenir une grande diversité de métadonnées, notamment de type Exif. Par contre, les fichiers PNG et GIF en contiennent généralement peu. En outre, si l'image a été publiée sur Facebook ou Instagram, de nombreuses métadonnées auront été supprimées par le système de gestion d'images de ces plateformes⁴. Les environnements à risque sont donc plutôt les forums, les blogues et les sites d'actualités, mais les paramètres de gestion des métadonnées dans les réseaux sociaux

3. Selon Friedland et Sommer (2010), seulement 4,3 % de photographies et de vidéos sur Flickr contiennent des balises de géolocalisation.

4. Les essais effectués pour l'écriture de cet article étaient limités à Facebook et à Instagram, mais d'autres contextes sensibles comme les applications de rencontre Tinder et Grindr mériteraient également d'être expérimentés.

FIGURE 14.2

À gauche, une image JPEG publiée en ligne, recadrée pour anonymiser le portrait. À droite, la vignette intégrée extraite du fichier JPEG, révélant l'image d'origine avant la manipulation (à partir de Murdoch et Dornseif, 2004)



peuvent être amenés à évoluer dans le futur.

La responsabilité épistémique dans le régime de visibilité numérique

Les possibilités de collecte et d'exploitation de métadonnées générées par l'activité des internautes amènent à interroger sous un nouveau jour la question de la responsabilité épistémique dans les activités de recherche. La notion de responsabilité épistémique est entendue ici comme l'ensemble des obligations liées à la production de connaissances. Développée dans le champ de l'épistémologie sociale, elle a été par la suite étendue à l'éthique de la recherche en contexte numérique, en s'appuyant notamment sur les théories féministes de la technologie ainsi que leurs réflexions sur le rapport entre savoir et pouvoir. Ainsi, la philosophe des sciences Judith Simon suggère que dans un contexte sociotechnique caractérisé par l'hyperconnexion, cette responsabilité doit être envisagée comme une charge distribuée entre plusieurs parties prenantes : elle engage non seu-

lement les chercheurs et chercheuses sur le plan individuel, mais elle implique aussi une réflexion sur la gouvernance des plateformes numériques, de façon à concevoir des systèmes et environnements permettant d'agir de manière responsable (Simon, 2015).

C'est d'autant plus le cas lorsque les méthodes employées mobilisent des scripts de collecte de contenu à grande échelle (*Web scrapping*) de façon à obtenir de larges échantillons d'images publiées sur des comptes personnels ou dans des groupes d'utilisateurs et d'utilisatrices. McQuire souligne que le principal droit réclamé par les sites de partage d'images est « le droit d'explorer les images (y compris leurs métadonnées) ainsi que les transactions des internautes qui postent et échangent des images pour en tirer des données [*data mining*] », de façon à modéliser « les réseaux de sociabilité dans lesquels ils circulent » (traduction libre, 2013, p. 237). Les nouvelles capacités d'action des équipes de recherche, renforcées par l'appui de nouveaux agents computationnels (robots, logiciels de traitement de l'information), reconfigurent le rapport entre le consentement des acteurs à partager des informations et le pouvoir des chercheurs et chercheuses à s'en saisir.

La capture d'images publiées sur le Web et l'exploitation de leurs métadonnées entrent dans la catégorie des méthodes dites discrètes ou non réactives, c'est-à-dire des méthodes de collecte indirectes ou dissimulées, impliquant des personnes qui ne savent pas qu'elles font l'objet d'une recherche. Lee (1993) distingue plus précisément les « données trouvées » (directement accessibles dans l'espace public) et les « données extraites » (tirées d'enregistrements ou d'archives). À cet égard, la condition des métadonnées intégrées est ambiguë : si une photographie est accessible publiquement sur le Web (donnée trouvée), ses métadonnées intégrées sont quant à elles des traces qui doivent être extraites du fichier (donnée extraite). Ainsi, pour un même fichier image, le statut des métadonnées intégrées et son rapport au consentement des internautes à partager ces données doivent être évalués de façon distincte par rapport à ceux de l'image affichée à l'écran.

Cette réflexion sur les dimensions éthiques de l'analyse des traces ne date toutefois pas de l'avènement du Web. Il est intéressant de constater qu'une réflexion similaire anime les débats sur la collecte et l'analyse du contenu des poubelles. Tout comme les ordures, les métadonnées « peuvent révéler des aspects de la vie de quelqu'un, qu'il ou elle préférerait garder

cachés» (traduction libre de Lee, 2000, p. 29). D'un point de vue légal, lorsqu'un déchet est disposé sur la voie publique, il pourrait raisonnablement être considéré comme abandonné. Cependant, Rathje et Cullen (1992) soulignent qu'aux États-Unis, «l'examen des ordures d'autrui constitue potentiellement une invasion de la vie privée, étant donné que les citoyens sont protégés contre "la recherche et la saisie déraisonnables" par le quatrième amendement» (traduction libre de Lee, 2000, p. 29). Dans une étude sur les «données sales» et les méthodes d'investigation des détectives privés, Shulman a proposé l'expression «vulnérabilité performative» pour décrire les conséquences d'actions qui dévoilent des informations compromettantes pour les personnes concernées, par exemple en rendant empiriquement accessibles des indices de tromperie ou de crime (Shulman, 1994, p. 231).

Pour diverses études s'intéressant principalement à l'analyse du contenu pictural des images et non de leurs métadonnées, les matériaux visuels collectés en ligne ne devraient pas être considérés comme des traces d'activité laissées par des sujets humains dont il faudrait avant tout protéger le droit à la vie privée, mais plutôt comme des documents publiés sur le Web – au même titre qu'une photographie ou un texte peuvent être publiés sur un support papier, par des personnes qui souhaitent que leur travail soit rendu public. Bassett et O'Riordan soutiennent ainsi que, dans le cas des textes publiés, les auteurs et autrices «prennent une décision sur leur niveau de visibilité, en contrôlant le degré de divulgation de leur identité» (traduction libre, 2002, p. 236). Cette approche est valable dans de nombreux contextes où les images sont produites lors d'une activité professionnelle, semi-professionnelle ou même amateur (par exemple, un blogue de voyage destiné à un large public).

Dans d'autres cas, lorsque l'activité de production visuelle s'inscrit dans une démarche d'autodocumentation de la vie quotidienne, générant des échanges manifestement restreints au cercle des connaissances personnelles, par exemple sur un compte Instagram ou sur un profil Facebook, même si cette activité est faite en public, il convient de poser un diagnostic préalable à la collecte de matériaux afin d'essayer de comprendre comment ces personnes négocient la frontière entre vie privée et vie publique, et dans quelle mesure la démarche méthodologique peut s'efforcer de respecter ces limites. En fin de compte, pour utiliser des méthodes dites discrètes en vue de collecter et d'exploiter des traces

numériques, il faut comprendre ces données comme des « productions socialement situées » prenant leur sens dans un contexte social, organisationnel et politique particulier (traduction libre de Lee, 2000, p. 85).

La visibilité contextuelle des métadonnées

En fin de compte, pour évaluer le consentement éclairé d'un internaute à partager des informations contenues dans les métadonnées intégrées, il est important de resituer leur degré de visibilité dans le contexte de la plateforme où l'image a été initialement publiée. On parlera de « visibilité contextuelle » pour décrire le degré variable de visibilité des métadonnées intégrées d'une image selon la plateforme où elle est affichée. Par exemple, dans l'interface de Flickr, lorsqu'une balise de géolocalisation est présente dans un fichier image, l'interface signale sa présence en pointant la localisation sur une carte géographique affichée à l'écran. Étant donné le caractère immédiatement visible de cette information sur l'interface, on peut supposer que la personne qui a publié l'image était consciente que cette information a été publiquement partagée lorsque l'image a été mise en ligne. Par contre, dans le cas du site Web de *Vice*, l'interface n'a pas été programmée pour afficher automatiquement les balises de géolocalisation des images; leur présence dans la photo de McAfee n'a donc pas été détectée lors de sa publication, entraînant son partage involontaire. La divulgation de cette information a ainsi changé son degré de publicité, qui passe de confidentiel à publiquement exposé (Latzko-Toth et Pastinelli, 2014).

Cela dit, il serait problématique d'associer systématiquement la visibilité des métadonnées affichées à l'écran et le consentement de l'internaute à partager des informations. En effet, d'autres paramètres entrent en ligne de compte pour évaluer les dimensions contextuelles de la vie privée en ligne (Markham et Buchanan, 2012). La notion de *privacy in context*, proposée par Nissenbaum (2009), vise à évaluer le degré de confidentialité d'une information. Elle donne au principe d'« intégrité contextuelle » une place centrale dans l'éthique de la recherche, c'est-à-dire que la collecte et la diffusion d'une information à des fins de recherche doivent tenir compte des normes et des paramètres initiaux définis lorsque l'information a été originalement publiée. En ce qui concerne les métadonnées intégrées aux images, ce principe d'intégrité contextuelle doit être arrimé aux dimensions clés de la visibilité et de l'invisibilité, qui elles aussi

dépendent du contexte des plateformes et des paramètres d'affichage des interfaces.

* * *

Les métadonnées intégrées sont à la fois des opérateurs de circulation des images et des microarchives mobiles qui enregistrent des traces d'activités humaines et computationnelles. Le présent chapitre a proposé de les analyser sous l'angle d'une approche forensique pour évaluer leur potentiel et leurs limites méthodologiques, ainsi que les défis éthiques soulevés. Ces corpus de traces témoignent des modes de production, de circulation et de consommation d'artefacts numériques. En tant que tels, ils ont pour effet de perturber la relation entre le consentement de l'internaute à partager des données sensibles et les indices d'activité disponibles en ligne. À mesure qu'ils circulent en ligne, les fichiers image sont soumis à une dynamique complexe d'accumulation et de dégradation des métadonnées, étroitement liée à l'infrastructure technique des plateformes. Ainsi, examiner le fonctionnement (et les dysfonctionnements) des métadonnées intégrées permet de mieux comprendre comment la circulation des images numériques est conditionnée par des rapports complexes de pouvoir économique et politique affectant le régime contemporain de visualité numérique. Dans un avenir proche, l'évolution des moyens techniques de distribution des images telles que les *iframes* (systèmes d'affichage centralisant la collecte de métadonnées transactionnelles) et les systèmes *blockchain* ou chaîne de blocs (notamment les *smart contracts*) pourraient offrir de nouvelles possibilités et poser de nouveaux défis pour l'analyse visuelle du Web.

Références

- BASSETT, Elizabeth H. et Kate O'RIORDAN, « Ethics of Internet research: Contesting the human subjects research model », *Ethics and Information Technology*, vol. 4, n° 3, 2002, p. 233-247.
- BRÜGGER, Niels, *The Archived Web: Doing History in the Digital Age*, The MIT Press, 2018.
- CASEMAJOR, Nathalie, « Digital materialisms: Frameworks for digital media studies », *Westminster Papers in Communication and Culture (WPCC)*, vol. 10, n° 1, 2015.

- FRIEDLAND, Gerald et Robin SOMMER, « Cybercasing the joint: On the privacy implications of geotagging », dans *Proceedings of the Fifth USENIX Workshop on Hot Topics in Security (HotSec 10)*, 2010, Washington, 6 p.
- GOODIN, Dan, « Feds charge confessed Anon member after tracking his digital footprints », *Ars Technica*, 4 avril 2012.
- HOCHMAN, Nadav et Lev MANOVICH, « Zooming into an Instagram City: Reading the local through social media », *First Monday*, vol. 18, n° 7, 2013.
- KIRSCHENBAUM, Matthew, *Mechanisms: New Media and the Forensic Imagination*, The MIT Press, 2008.
- LATZKO-TOTH, Guillaume et Madeleine PASTINELLI, « Par-delà la dichotomie public/privé: la mise en visibilité des pratiques numériques et ses enjeux éthiques », *tic&société*, vol. 7, n° 2, 2014.
- LEE, Raymond M., *Unobtrusive Methods in Social Research*, Open University Press, 2000.
- LEE, Raymond M., *Doing Research on Sensitive Topics*, SAGE Publications, 1993.
- MARKHAM, Annette N. et Elizabeth BUCHANAN, *Ethical Decision-Making and Internet Research: Recommendations from the AoIR Ethics Working Committee (Version 2.0)*, Association of Internet Researchers, 2012.
- MCQUIRE, Scott, « Photography's afterlife: Documentary images and the operational archive », *Journal of Material Culture*, vol. 18, n° 3, 2013, p. 223-241.
- MURDOCH, Steven J. et Maximilian DORNSEIF, « Far more than you ever wanted to tell: Hidden data in Internet published documents », dans *21st Chaos Communication Congress*, 2004, p. 27-29.
- NISSENBAUM, Helen, *Privacy in Context: Technology, Policy, and the Integrity of Social Life*, Stanford University Press, 2009.
- RATHJE, William L. et Cullen MURPHY, *Rubbish! The Archaeology of Garbage*, Harper Collins, 1992.
- SIMON, Judith, « Distributed epistemic responsibility in a hyperconnected era », dans Luciano FLORIDI (dir.), *The Onlife Manifesto*, Springer, 2015, p. 145-159.
- SHULMAN, David, « Dirty data and investigative methods: some lessons from private detective work », *Journal of Contemporary Ethnography*, vol. 23, n° 2, 1994, p. 214-253.

CHAPITRE 15

Les contextes d'usage de la téléphonie mobile : données fines et enregistrement vidéo

Julien Figeac, Johann Chaulet

Les développements récents des téléphones dits intelligents et la diffusion de l'accès à Internet en situation de mobilité ont profondément diversifié le champ des pratiques de la téléphonie mobile. Près de 73 % de la population française est dorénavant détentrice d'un téléphone mobile connecté (CREDOC, 2017). Il est difficile de proposer un panorama exhaustif des nombreuses répercussions sociales et culturelles engendrées par la démocratisation progressive de l'accès aux multiples applications commercialisées au sein des plateformes des principaux systèmes d'exploitation (Android, Apple ou Windows Phone). Il faut toutefois rappeler que les individus se détournent progressivement de leurs ordinateurs et préfèrent dorénavant fréquenter les médias sociaux depuis leurs terminaux mobiles (CREDOC, 2017). Les travaux traitant spécifiquement des usages nomades des réseaux socionumériques, tels que Facebook ou Twitter, sont encore rares en sciences humaines et sociales. Pourtant l'enjeu est de taille tant cette appropriation de la téléphonie mobile forme un creuset technologique central dans la configuration contemporaine des sociétés et de l'individualisme en réseaux (Rainie et Wellman, 2012).

Si la compréhension de ces pratiques relationnelles médiatisées et de plus en plus mobiles constitue un important enjeu scientifique, il est également essentiel de mieux comprendre comment la palette d'applications

des téléphones mobiles renouvelle les pratiques culturelles et médiatiques. Cette contribution méthodologique part de l'idée que la carence des recherches actuelles dans ce domaine tient en partie aux difficultés méthodologiques soulevées quand on entreprend de documenter, de manière fine et rigoureuse, ce type d'usage. Pour aller en ce sens, il s'agira ici de présenter une méthode de recherche fondée sur un protocole d'enregistrements audiovisuels, qui permet de collecter des données détaillées et robustes par rapport à ces utilisations contemporaines de la téléphonie mobile (Licoppe et Figeac, 2018).

Le protocole en question a été développé pour collecter des traces d'usage des téléphones mobiles à l'aide d'enregistrements audiovisuels. Ces données vidéo se caractérisent par la synchronisation de deux enregistrements complémentaires : l'enregistrement des activités réalisées depuis les écrans tactiles des téléphones mobiles à l'aide d'une application spécifique, et l'enregistrement vidéo des contextes d'usage à l'aide de lunettes caméra portées par les individus. La synchronisation de ces deux prises de vue, sur les écrans et les contextes d'usage, permet de recueillir des données particulièrement fines sur la façon dont les individus gèrent leur engagement conjoint dans l'utilisation du téléphone mobile et dans la conduite de leurs déplacements ordinaires.

Les défis méthodologiques soulevés par l'étude des usages des téléphones en mobilité peuvent être abordés successivement, soit par le recueil de données contextuelles sur les usages des téléphones mobiles grâce à l'utilisation de lunettes caméra, le recueil de *patterns* d'usage des téléphones mobiles grâce à l'enregistrement de l'activité d'écran, et l'analyse des usages par la synchronisation des traces d'utilisation et des données contextuelles. Tenir compte de ces défis permet d'envisager l'importance des entretiens face aux implications éthiques de ce type de méthodes vidéo-ethnographiques pour l'étude des usages de la téléphonie mobile.

Les défis méthodologiques

L'appropriation des téléphones mobiles et de leurs applications exploitant les ressources d'Internet mobile a été étudiée en recourant à diverses méthodes, notamment grâce au recueil et au traitement statistique de journaux de connexion (Tossell *et al.*, 2012). Malgré leur intérêt propre, ces traces d'usage ne permettent pas de documenter convenablement les

situations d'utilisation réelles, si ce n'est par des inférences indirectes. D'autres approches plus qualitatives (Isomorsu *et al.*, 2007), fondées sur des entretiens ou sur la méthode des carnets d'utilisation, rendent compte de manière plus fine de l'expérience d'utilisation. Malgré leur granularité d'analyse relativement fine, ces approches ne permettent pas d'articuler concrètement le déroulement des activités technologiques avec les situations de mobilité. Pour atteindre cet objectif, il est apparu pertinent de concevoir un dispositif d'observation fondé sur des enregistrements vidéo.

De nombreuses recherches ont permis de créer ce type de protocoles méthodologiques fondés sur des données audiovisuelles pour analyser de manière détaillée les comportements et les activités au sein de diverses organisations sociales, notamment les modes de collaboration dans les espaces de travail (Lahlou, 2011), ainsi que diverses pratiques relevant des champs de la culture ou de la communication (Charbey *et al.*, 2017). Ces protocoles de recherche audiovisuels, fondés sur des caméras fixes installées à des endroits précis par rapport aux situations à analyser, ont été formalisés avec les années au point que la mise en place d'une telle méthode est aujourd'hui bien balisée (Heath *et al.*, 2007). Des travaux ont permis d'appliquer cette orientation méthodologique à l'étude de la téléphonie : ils ont commencé par utiliser des caméscopes pour accéder au plus près à l'expérience des mobinautes en enregistrant leurs usages, en récoltant des données sur leur manière de gérer les applications mobiles de leur téléphone, tout en étant sensibles aux contextes et aux autres activités réalisées en parallèle de leur utilisation. Cependant, les caméscopes ne permettent pas de suivre les individus durant leurs déplacements quotidiens pour réellement comprendre comment ces technologies de poche sont utilisées en situation de mobilité.

Pour mettre en place une telle méthode, il est pertinent de s'inspirer de l'anthropologie visuelle qui a étudié la mobilité humaine à l'aide de caméras portatives. Cela permet notamment de filmer l'itinéraire de l'individu tout en enregistrant ses propres commentaires pour produire une visite visuelle égocentrée de différents environnements urbains (Oulasvirta *et al.*, 2005). Cette captation audiovisuelle, réalisée à la première personne, a également été développée dans le cadre d'approches ethnographiques des usages nomades des téléphones mobiles (Zouinar *et al.*, 2004). Dans un tel cas, les caméras portatives sont surtout utilisées pour documenter la multimodalité des interactions sociales en ciblant les

contextes de mobilité où il est facile de « suivre » les individus, par exemple en équipant les automobiles de caméras (Laurier *et al.*, 2008). Certains de ces travaux montrent ainsi comment les personnes gèrent de façon collaborative, en situation de conduite, les événements et les perturbations liés aux téléphones mobiles (Haddington et Rauniomaa, 2011). Cependant, les protocoles vidéo déployés dans les habitacles des véhicules ne sont pas adaptés aux autres situations de mobilité, par exemple aux transports en commun. L'enjeu est pourtant majeur pour comprendre la diversité des utilisations qui sont effectuées. Pour mettre en place un protocole de recherche robuste dans ces circonstances denses et complexes, nous proposons une méthode qui exploite des dispositifs d'enregistrement vidéo portatifs susceptibles d'être « portés » par les individus.

La cueillette de données contextuelles avec des lunettes caméra

Le choix d'un tel dispositif portatif va être fonction du type de données empiriques recherchées et va orienter le déroulement de l'analyse. Les caméras fixées sur le torse de la personne, à l'aide d'un harnais par exemple (voir la figure 15.1), permettent de capter une prise de vue contextuelle qui fournit des indications sur ses postures et ses orientations corporelles. Elles donnent par exemple la possibilité d'étudier comment une joueuse se coordonne avec son équipe lors d'une partie de jeu vidéo en réseau fondé sur les ressources de la géolocalisation (voir la figure 15.2). Il devient alors possible d'entendre toutes les interactions verbales des personnes participantes et de saisir certaines de leurs expressions déictiques afin de décrire comment elles s'approprient la carte du jeu ainsi que la manière dont leurs adversaires ou les objets des quêtes sont représentés.

Il est indiqué de mobiliser ce type de dispositif portatif quand il n'est pas possible de suivre les individus en situation (Charbey *et al.*, 2017) pour que les caméras puissent collecter des traces des pratiques. Grâce à ces traces, il devient possible de réaliser des entretiens *a posteriori* en les faisant visionner aux participants. L'objectif est alors de les amener à rendre compte de leurs pratiques, de la manière dont ils ont utilisé un service mobile, des problèmes rencontrés, etc. Ce niveau de collecte des données est bien souvent pertinent dans l'optique des recherches en sociologie de la culture, par exemple, pour documenter des phénomènes comme l'appropriation des applications mobiles développées dans le cadre d'un dispositif

FIGURE 15.1

Caméra fixée sur un harnais

FIGURE 15.2

Équipe d'un jeu géolocalisé en action

de médiation muséal multimédia. La prise de vue des enregistrements vidéo ne donne pas réellement accès à l'expérience de l'individu, mais permet de percevoir ce qu'il fait plus concrètement avec le dispositif technologique. Pour accéder à cette expérience, on peut réaliser l'enregistrement vidéo en positionnant la caméra au niveau de la tête de la personne, car cela permet de donner accès à ce qu'elle voit et de le reproduire. Cette captation vidéo, du point de vue de la personne et de son engagement visuel dans l'action ou « en action », est facilitée par la miniaturisation des caméras portatives qui sont dorénavant relativement discrètes. De même, le perfectionnement des oculomètres permet aujourd'hui à la psychologie

d'affiner l'analyse des processus cognitifs engagés dans les usages des téléphones mobiles (Drewes *et al.*, 2007). Il apparaît malgré tout que les oculomètres et leurs protocoles expérimentaux restent difficilement transposables aux contextes réels de la mobilité humaine. Ces contraintes ont justifié le développement d'approches plus naturalistes fondées sur le fait de demander aux personnes de porter une caméra frontale ou des lunettes spécialement équipées d'une microcaméra et d'un microphone ; elles sont désignées de différentes manières au sein de la littérature scientifique : « lunettes vidéo », « *subcams* » ou « lunettes caméra » (voir la figure 15.3). Ces lunettes ont été conçues pour capter le point de vue subjectif d'un sujet engagé dans une activité sociale, notamment pour l'amener à visionner les enregistrements afin qu'il puisse rendre compte, de manière rétrospective, du vécu de son expérience (Lahlou, 2011). Ces lunettes ont ensuite été utilisées pour étudier les usages des téléphones mobiles en situation naturelle (Relieu, 2009), afin de fournir notamment des informations contextuelles susceptibles d'enrichir le traitement quantitatif de traces d'usage des services mobiles (Zouinar *et al.*, 2004).

Ces enregistrements vidéo peuvent être traités de manière autonome pour étudier, par exemple, les usages des téléphones mobiles, d'Internet et de la télévision dans les transports publics (voir la figure 15.4 ; Figeac, 2009). Ils donnent ainsi la possibilité d'étudier les contextes d'utilisation et la manière dont ils influent sur les usages. Il est alors possible de montrer que les individus privilégient, quand ils sont dans les transports souterrains, les applications qui ne demandent pas un accès fiable et permanent au réseau téléphonique, comme la lecture des fils d'actualités, plutôt que les services qui sont plus sensibles aux interruptions de flux, comme la consultation de vidéos. Ces enregistrements permettent également de comprendre comment la promiscuité ou l'impossibilité de trouver une place assise rendent l'usage de certains services plus pertinent que d'autres, notamment parce qu'ils nécessitent un moindre degré d'attention. Cela étant dit, le champ des lunettes caméra étant très restreint, l'enregistrement vidéo qu'elles génèrent ne procure qu'un accès partiel au déroulement de ce qui se passe à l'écran (voir la figure 15.4). Il est donc utile d'essayer de compléter cette prise de vue contextuelle réalisée avec les lunettes par l'enregistrement de la capture de l'écran pour avoir accès de façon précise et continue aux activités que les individus réalisent à l'aide de leur mobile.

FIGURE 15.3

Lunettes caméra



FIGURE 15.4

Vidéo réalisée avec des lunettes caméra



La documentation des patterns d'usage avec l'enregistrement de l'activité d'écran

Pour pouvoir rendre compte de l'expérience des individus et du déroulement de leurs pratiques, il est nécessaire d'avoir également accès à ce qui est affiché sur leurs écrans. Il était jusqu'à récemment très difficile d'enregistrer les usages des téléphones mobiles à même l'appareil : cela se faisait à l'aide d'un enregistrement vidéo externe (Morel et Licoppe, 2011).

Dorénavant, les capacités graphiques et de calcul des téléphones mobiles permettent de réaliser des captures de l'activité d'écran avec une application dédiée installée directement sur l'appareil (Figeac et Chaulet, 2018). Il existe plusieurs applications qui remplissent cette fonction. Pour ne pas être contraint de déverrouiller le terminal afin de pouvoir bénéficier des droits d'un « super-usager », il est préférable d'en privilégier une qui puisse fonctionner en arrière-plan des différentes tâches du téléphone (par exemple Screen Recorder). Ce type d'application ne fonctionne qu'avec le système d'exploitation Android ou Windows mobile, car Apple n'autorise pas leur usage. Notre protocole méthodologique ne peut donc pas être utilisé, en ce moment du moins, pour étudier les usages des iPhones, autrement dit les utilisations de 14,6 % des téléphones intelligents commercialisés aujourd'hui dans le monde (contre 86,2 % des téléphones qui fonctionnent avec Android) (Gartner, 2016).

Ces applications enregistrent l'écran et l'ensemble des actions réalisées à l'aide de l'écran tactile du téléphone. Ce dispositif fournit ainsi des données très précieuses et particulièrement détaillées sur les usages des services mobiles. Il permet de collecter et de décrire la manière dont l'individu navigue au sein de son téléphone, comment il utilise le *design* des services mobiles ainsi que leurs diverses fonctionnalités. Plus généralement, il offre la possibilité d'analyser le déroulement et l'organisation temporelle des *patterns* d'usage des téléphones mobiles en mesurant la fréquence et la durée d'utilisation des différentes applications. Il devient ainsi possible de montrer comment les applications des réseaux sociaux numériques, et plus particulièrement de Facebook, occupent une place centrale dans les utilisations. En effet, grâce à la prégnance de leurs dispositifs de notification qui visibilisent en continu différents événements importants tels que la réception des messages entrants (voir les figures 15.5 à 15.8), ces services incitent les usagers à traiter en priorité le contenu de ces alertes. Il apparaît alors que les sollicitations relationnelles médiatisées par les interfaces des téléphones incitent les individus à adopter des pratiques de connexion permanentes (Figeac et Chaulet, 2018).

Ces images illustrent la diversité des usages qui peuvent être étudiés à l'aide de cette méthode. En donnant accès à ce qui est affiché sur les écrans, notamment au contenu des messages, les captures d'écran offrent une réelle richesse d'analyse et permettent d'étudier le détail des pratiques réalisées depuis les téléphones mobiles. Nous défendons cependant l'idée

FIGURE 15.5

Traitement des modifications

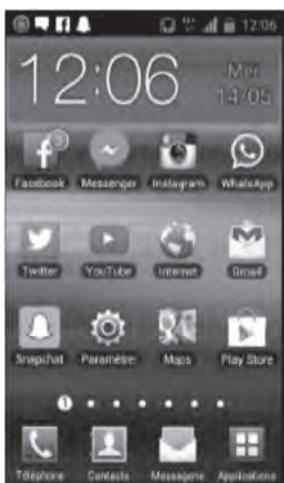


FIGURE 15.6

Traitement des sollicitations sociales



FIGURE 15.7

Priorisation d'une relation



FIGURE 15.8

Lecture du message privé



qu'il est préférable d'articuler ce niveau d'analyse avec l'analyse de données contextuelles plus précises. Pour collecter ce type de données, il convient de demander aux individus de filmer leurs usages à l'aide de lunettes caméra. Il devient ainsi possible de synchroniser les deux prises de vue dont il a été question ici, la captation visuelle de l'activité d'écran à l'aide d'une application dédiée installée sur les téléphones et l'enregistrement vidéo des contextes d'usage recueilli à l'aide des lunettes caméra (voir la figure 15.9).

FIGURE 15.9

Synchronisation de l'enregistrement de l'utilisation et de la prise de vue contextuelle



L'analyse des usages par les traces d'utilisation et les données contextuelles

La synchronisation de ces deux enregistrements permet de décrire le déroulement des utilisations en détail et la manière dont elles sont organisées en contexte. Dans les situations de mobilité, les usages des téléphones mobiles sont étroitement articulés avec la conduite des déplacements ordinaires, la gestion des environnements et la coordination avec les autres individus. Comme les individus greffent les activités réalisées depuis leurs téléphones mobiles sur le déroulement de l'activité de déplacement, ils s'engagent dans des formes de multi-activité complexes de plus en plus organisées autour des téléphones mobiles et de leur *design*. Ce protocole méthodologique permet d'analyser finement la configuration de ces multi-activités (Licoppe et Figeac, 2018).

Ces données audiovisuelles permettent ainsi de décrire comment les personnes se déplacent et se coordonnent dans les espaces publics et comment les téléphones mobiles s'inscrivent dans le cours des actions des individus en situation de mobilité (Figeac, 2009). Elles mettent à la fois en lumière la manière dont les utilisations des dispositifs technologiques sont négociées et ajustées aux contraintes d'usage qui émergent des situations de mobilité ordinaires, en montrant par exemple comment le choix des applications (SMS, jeux, presse, médias sociaux, etc.) est lié à la configuration des situations vécues et aux postures corporelles des individus. Elles révèlent réciproquement comment les téléphones mobiles orientent les pratiques sociales des individus, leurs déplacements et leurs modes de participation dans les espaces publics. Par exemple, ces traces d'usage audiovisuelles mettent en évidence la manière dont les médias sociaux incitent les usagers et usagères à répondre rapidement aux messages reçus par l'affichage continu de notifications et viennent renouveler l'économie des échanges de regards dans les lieux publics. Les notifications invitent en effet les individus à leur accorder une attention latente et à multiplier les coups d'œil pour vérifier la réception de nouveaux messages. Cette incitation technologique à maintenir une connexion permanente avec ses cercles relationnels vient modifier la manière dont les individus occupent les espaces d'une part, et elle renouvelle l'économie des formes visuelles de la coordination sociale dans les environnements urbains, d'autre part (Figeac et Chaulet, 2018).

Les implications éthiques des entretiens

Dans la mesure où toutes les méthodes d'enquête sur les usages de la téléphonie mobile demandent de négocier au mieux le caractère intrusif de la collecte de données, que ce soit à l'aide de sondes ou depuis les serveurs des opérateurs, on doit en considérer les implications éthiques. Le protocole audiovisuel décrit ici soulève moins d'enjeux éthiques que d'autres méthodes, car les données personnelles qui sont directement affichées sur les écrans sont moins nombreuses, notamment lorsque les individus jouent, lisent des contenus ou regardent des clips et des vidéos. Les enregistrements vidéo capturent cependant bon nombre de données personnelles particulièrement sensibles, comme le contenu de certains SMS ou certains commentaires publiés dans les médias sociaux. Lors de la consultation de ces

applications, toutes les informations affichées à l'écran sont effectivement recueillies (le nom des personnes, leur photo, etc.). Il est alors nécessaire de mettre en place un ensemble de précautions pour éviter de porter atteinte à la vie privée des individus et de ceux avec qui ils partagent des informations à l'aide de leur téléphone.

Il convient tout d'abord de considérer qu'avec ce protocole de recherche, c'est la personne faisant usage du téléphone mobile qui gère la collecte des données. Elle active l'enregistrement de son téléphone quand elle le souhaite et peut l'éteindre à tout moment. Le logo représentant une petite caméra, dans la barre de notification située dans le coin supérieur de l'écran, lui rappelle qu'elle est en train d'enregistrer ses usages et que tout ce qu'elle consulte va être mémorisé dans la vidéo. Elle peut également visionner les enregistrements à même son téléphone. Si elle le souhaite, elle peut supprimer ceux qu'elle juge trop intrusifs par rapport à sa vie privée. Cela étant dit, il faut garder en tête que les informations ou les contenus publiés dans les médias sociaux sont bien souvent publiés avec l'idée d'être partagés auprès d'un large nombre de destinataires. Il est donc très rare que les individus jugent ces messages ou ces contenus trop sensibles et intrusifs; nous n'avons pour notre part jamais rencontré de telles suppressions d'enregistrements dans nos projets de recherche.

Même si ces précautions permettent de donner aux personnes participantes des moyens pour filtrer leurs données, il en va de la responsabilité de chacun d'en faire un usage respectueux. Nous considérons que la simple signature d'un accord de consentement n'est pas suffisante et pose problème sur le plan éthique, même si elle est acceptée légalement. La raison en est simple. Tout d'abord, les personnes n'ont pas forcément conscience de la nature des données collectées, du moins chez celles qui n'auraient pas eu la curiosité de visionner les enregistrements vidéo réalisés. De plus, il paraît important que le chercheur ou la chercheuse leur fasse part de son point de vue, en leur expliquant comment ces données audiovisuelles seront traitées et quels phénomènes seront explorés tout en recueillant leurs avis et en tenant compte de leur point de vue et de la manière dont elles expliquent leurs usages et pratiques. C'est pourquoi nous préconisons de réaliser des entretiens afin de leur montrer des séquences vidéo et de discuter avec eux de leur contenu.

Le but de ces entretiens est de demander aux individus d'expliquer leurs usages, de décrire les procédures pratiques qu'ils mettent en œuvre

pour utiliser les services mobiles, et pour s'approprier les interfaces, notamment (voir les chapitres de Gallant *et al.* et Pélissier, le présent ouvrage). Il est intéressant de solliciter leur point de vue, par exemple en leur faisant relire leurs échanges dans les médias sociaux, notamment ceux que l'on souhaite étudier et mobiliser comme des *verbatim* lors de la rédaction des articles scientifiques. En plus d'être une manière de les impliquer et de leur faire prendre la mesure du traitement de leurs données, ces entretiens viennent enrichir l'analyse des traces en permettant de recueillir des informations précieuses sur la nature du lien entre les individus et sur le contexte d'énonciation de leurs messages. On peut ainsi mieux comprendre notamment l'historique des échanges médiatisés et leur arrière-plan personnel, affectif et relationnel.

Ces entretiens ont dès lors une finalité éthique, car il s'agit de montrer et d'expliquer aux individus l'analyse qui va être faite de leurs enregistrements vidéo. Et ils ont une finalité scientifique dans la mesure où il s'agit d'impliquer ces individus dans l'analyse proprement dite. Seuls des entretiens permettent de collecter des récits susceptibles d'inscrire les utilisations capturées par les vidéos dans le temps long des pratiques sociales. C'est ainsi qu'il est par exemple possible de dire que les dispositifs de notification des médias sociaux favorisent la connexion permanente, quand les individus constatent par eux-mêmes ce phénomène et expliquent la pression relationnelle grandissante dont ils font l'objet.

Malgré ce positionnement scientifique et éthique, d'autres problèmes subsistent. Ces précautions ne permettent pas d'informer tous les correspondants téléphoniques ou tous les « amis » des personnes participantes de l'usage qui est fait des informations qu'ils ont pu échanger avec eux. Ce problème n'est pas nouveau en sociologie dans la mesure où toutes les recherches sur les pratiques téléphoniques et sur les réseaux de relations sociales sont fondées sur la collecte d'informations sociodémographiques intimes fournies par les personnes participantes pour documenter l'appartenance sociale et les conditions de vie de leurs relations. Dans le cas des protocoles audiovisuels, il est possible de procéder à un traitement anonyme des données en floutant le nom, les photos et toutes les informations personnelles qui ont été collectées. Il convient également d'éviter de publier des captures d'écran qui montrent des personnes, filmées par exemple dans des environnements publics, ou de flouter leurs visages le cas échéant (voir la figure 15.4).

Finalement, le protocole de recherche audiovisuel fondé sur la collecte et la synchronisation de deux prises de vue permet la capture des activités réalisées depuis les écrans tactiles des téléphones et l'enregistrement des contextes d'usage à l'aide de lunettes caméra. Ces enregistrements sont particulièrement pertinents, car ils fournissent des données très précises sur les logiques d'utilisation des différentes applications mobiles ainsi que sur les contextes d'usage. Il est possible de faire une analyse autonome de ces données, en examinant par exemple le contenu des messages ou l'organisation des fils de discussion dans une messagerie instantanée. Et il est généralement davantage possible d'analyser la structuration temporelle des *patterns* d'usage, la prise en main et la durée d'utilisation de chaque application mobile en fonction des circonstances et des contraintes émergeant des situations ordinaires.

Nous préconisons cependant de réaliser des entretiens auprès des individus qui ont fait les enregistrements. Ces vidéos vont leur permettre d'expliquer concrètement leurs usages et la manière dont ils s'approprient leur téléphone. Ils pourront mettre en perspective ce que les captures d'écran donnent à voir en resituant la spécificité d'une utilisation dans le temps long de leurs pratiques.

Cette méthode articulant des enregistrements vidéo et des entretiens est fructueuse sur le plan scientifique car elle permet de solliciter la réflexivité des individus par rapport à la structuration de leurs pratiques, tout en favorisant des récits d'expérience ancrés dans l'ordinaire des utilisations. Sur le plan éthique, elle permet aussi de collecter des données intimes, parfois très intrusives, tout en impliquant les individus dans l'analyse qui en est faite.

Références

- CHARBEY, Raphaël *et al.*, « Triangulation des méthodes pour l'étude des mobilités d'un jeu géolocalisé (*Ingress*) », *Terminal. Technologie de l'information, culture & société*, n° 121, 2017.
- CREDOC, *Baromètre du numérique*, rapport réalisé pour le Conseil général de l'économie, de l'industrie, de l'énergie et des technologies (CGE), l'Autorité de Régulation des Communications Electroniques et des Postes (Arcep) et l'Agence du numérique, 2017.
- DREWES, Heiko, Alexander DE LUCA et Albrecht SCHMIDT, « Eye-gaze interaction for mobile phones », dans *Mobility, '07 Proceedings of the 4th International Conference on Mobile Technology, Applications, and Systems and the 1st Interna-*

- tional Symposium on Computer Human Interaction in Mobile Technology*, Singapour, 2007, p. 364-371.
- FIGEAC, Julien, « L'appropriation de la Mobile TV autour des réseaux de communication », *Réseaux*, n° 156, 2009, p. 81-111.
- FIGEAC, Julien et Johann CHAULET, « Video-ethnography of social media apps' connection cues in public settings », *Mobile Media and Communication*, vol. 6, n° 3, 2018, p. 407-427.
- GARTNER, « Gartner says five of Top 10 worldwide mobile phone vendors increased sales in second quarter of 2016 », communiqué de presse, Egham, 19 août 2016.
- HADDINGTON, Pentti et Mirka RAUNIOMAA, « Technologies, multitasking, and driving: Attending to and preparing for a mobile phone conversation in a car », *Human Communication Research*, vol. 37, n° 2, 2011, p. 223-254.
- HEATH, Christian, Paul LUFF et Marcus SANCHEZ SVENSSON, « Video and qualitative research: Analysing medical practice and interaction », *Medical Education*, vol. 41, n° 1, 2007, p. 109-116.
- ISOMORSU, Minna *et al.*, « Experimental evaluation of five methods for collecting emotions in field settings with mobile applications », *International Journal of Human-Computer Studies*, vol. 65, n° 4, 2007, p. 404-418.
- LAHLOU, Saadi, « How can we capture the subject's perspective? An evidence-based approach for the social scientist », *Social Science Information*, vol. 50, n°s 3-4, 2011, p. 607-655.
- LURIER, Éric *et al.*, « Driving and "passenger": Notes on the ordinary organization of car travel », *Mobilities*, vol. 3, n° 1, 2008, p. 1-23.
- LICOPPE, Christian et Julien FIGEAC, « Gaze patterns and the temporal organization of multiple activities in mobile smartphones uses », *Human-Computer Interaction*, vol. 33, n°s 5-6, 2018, p. 311-334.
- MOREL, Julien et Christian LICOPPE, « Studying mobile video telephony », dans Monica BÜSCHER, John URRY et Katian WITSCHGER (dir.), *Mobile Methods*, Routledge, 2011, p. 164-182.
- OULASVIRTA, Antti *et al.*, « Interaction in 4-second bursts: The fragmented nature of attentional resources in mobile HCI », dans *Proceedings of the SIGCHI Conference on Human factors in Computing Systems*, Portland, Oregon, 2005, p. 919-928.
- RAINIE, Lee et Barry WELLMAN, *Networked: The New Social Operating System*, The MIT Press, 2012.
- RELIEU, Marc, « Mobile phone "work": Disengaging and engaging mobile phone activities with concurrent activities », dans Rich LING et Scott W. CAMPBELL (dir.), *The Reconstruction of Space and Time: Mobile Communication Practices*, Transaction Publishers, 2009, p. 215-229.
- TOSSELL, Chad *et al.*, « Characterizing Web use on smartphones », dans *CHI '12 Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, Austin, Texas, 2012, p. 2769-2778.
- ZOUINAR, Moustapha *et al.*, « Observation and capture of multimodal interaction in mobile situations », dans *UbiMob, '04 Proceedings of the 1st French-Speaking Conference on Mobility and Ubiquity Computing*, Nice, 2004, p. 5-8.

CHAPITRE 16

La boîte noire de Netflix : la configuration des usages par les interfaces et les algorithmes

Gabrielle Silva Mota Drumond

La dissémination des logiciels et leur insertion dans les activités quotidiennes soulèvent de nombreuses discussions, d'une part autour de l'emploi d'algorithmes dans la gestion des informations (Beer, 2013; Manovich, 2013) et, d'autre part, autour de la prescription d'usages. Ces technologies sont présentes dans notre quotidien sous différentes formes. Les applications mobiles que nous utilisons tous les jours pour communiquer avec d'autres personnes et les plateformes qui permettent aux compagnies de gérer les ventes de leurs produits et de leurs services représentent des exemples de logiciels insérés dans notre routine. Constitutifs de ces logiciels, les algorithmes sont définis dans le *Larousse* (2018) comme un « ensemble de règles opératoires dont l'application permet de résoudre un problème énoncé au moyen d'un nombre fini d'opérations ». Ils fonctionnent comme des filtres de contenus et des organisateurs des informations (Sharma, 2013). Les algorithmes renferment un système de valeurs (Cardon, 2015) qui catégorise et hiérarchise les informations, les individus, les publics, etc. Dans le cas de logiciels comme celui du service de vidéo à la demande Netflix, les algorithmes assurent une fonction de décision qui n'est pas neutre dans la vie des gens, lorsqu'ils tentent de représenter les goûts des personnes en les réduisant à un système de valeurs numériques (Beer, 2013).

Ce chapitre propose un appareillage méthodologique ancré dans les théories de la conception et les *software studies*, permettant d'étudier l'« usager configuré » (Woolgar, 1990) par ces logiciels, à partir d'une étude de cas des algorithmes de recommandation de Netflix. Cet appareillage vise à permettre l'analyse de la place des usagers dans le développement de ce type de logiciel. À la suite d'une mise en contexte sur la prolifération des algorithmes de recommandation, le cas à l'étude et l'orientation théorique retenue sont exposés, suivis de la méthodologie mobilisée. Cette méthodologie est présentée en trois temps : l'étude des algorithmes ou « boîte noire »¹, l'étude de l'interface du logiciel et l'étude des plans de coordination des usages. En conclusion, nous abordons les possibilités d'application de cet appareillage méthodologique dans d'autres types d'études de cas.

Les algorithmes et les « décisions » de nos visionnements

En 2012, Netflix a annoncé un nouveau système de recommandation de contenus visant à classer son catalogue de vidéos selon l'usage de rangées titrées, comme les rangées *Top 10 for* « nom du profil » et *More like* « nom du contenu visionné ». En expliquant les raisons pour lesquelles un contenu a été recommandé, Netflix tente alors de favoriser la confiance des usagers en son système de recommandation et d'encourager des pratiques particulières de consommation. À ce propos, Netflix affirme que « *This not only promotes trust in the system, but encourages members to give feedback that will result in better recommendations*² » (Netflix Technology Blog, 2012). Par exemple, les marathons de visionnement sont considérés par le système comme des exemples positifs dans la création des recommandations. Autrement dit, lorsqu'un usager regarde des épisodes d'une série les uns à la suite des autres, le système considère que ce type de contenu est apprécié par cet usager. Cette pratique de *binge watching*, ou écoute en rafale, est actuellement très populaire chez les usagers et usagères de Netflix.

1. Burrell (2016) décrit la « couche cachée » des logiciels, celle où se développe le traitement des données, comme une « boîte noire ». En se penchant sur les algorithmes de *machine learning* (apprentissage automatique), elle souligne l'opacité des calculs comme une des caractéristiques de telles technologies (Drumond, 2016).

2. « Cela favorise non seulement la confiance dans le système, mais encourage les membres à donner des rétroactions qui se traduiront par de meilleures recommandations » (Netflix Technology Blog, 2012).

En utilisant des algorithmes, Netflix prétend être capable de déduire les comportements des gens et de créer des stratégies de recommandation personnalisées de contenus à partir de ces données. Le cas de Netflix révèle les capacités des systèmes de recommandation et du traitement des données effectués par des algorithmes. Il met aussi en lumière la relation dialogique entre ce qui est établi par les consignes définies par les personnes à l'origine de la conception de tels systèmes et l'usage effectif qu'en font les gens. Cette relation est médiée par l'interface du logiciel et modélisée par les instructions de conception. L'usage effectif émerge ainsi d'une négociation entre ce qui est proposé par Netflix (usage configuré) et ce que les gens désirent faire durant leur visionnement sur la plateforme.

En fonction des pratiques des usagères et usagers et des données recueillies, les systèmes de recommandation, en tant que technologies sous-tendues par des algorithmes, valorisent certains contenus au détriment d'autres (Sharma, 2013), ce qui définit les sélections et les classifications des informations (Gillespie, 2014). Des travaux en *softwares studies*, tels que ceux de Manovich (2013) et Striphas (2015), indiquent que l'emploi de ces filtres de contenu, insérés dans une culture algorithmique, contribue au bouleversement des pratiques quotidiennes. D'une part, les logiciels jouent un rôle majeur dans la popularisation des produits informatiques, c'est-à-dire dans l'usage généralisé de tels produits dans plusieurs secteurs de la vie de tous les jours. D'autre part, les résultats issus du traitement des données entraînent des modifications dans la culture contemporaine (Manovich, 2013).

Pour Kitchin et Dodge (2011), les algorithmes représentent des entités omniprésentes, bien qu'ils restent encore obscurs (Beer, 2013), et leurs effets sur les pratiques des gens sont encore peu explorés (Beer, 2013; Hallinan et Striphas, 2016). Dans le but de comprendre le rôle de ces objets techniques dans la culture contemporaine, ainsi que les forces sociales et culturelles qui contribuent à les construire, Manovich (2013) plaide pour la réalisation de recherches spécifiques sur ces technologies. Le processus de catégorisation et de sélection des informations, effectué selon des critères plus ou moins clairs au regard des usagers, renvoie également à des enjeux sociopolitiques. Comme le souligne Gillespie, « la catégorisation est une intervention sémantique et politique puissante: quelles sont les catégories, qu'est-ce qui appartient à une catégorie, et qui décide de la mise en œuvre de ces catégories dans la pratique? » (traduction libre, 2014, p. 171) sont autant de questions à se poser.

Derrière une promesse d'objectivité ou de neutralité, les algorithmes et les personnes qui les conçoivent jouent un rôle politique important, puisqu'ils classent et hiérarchisent à la fois les choses (idées, objets, informations, etc.) et les publics de façon préalable. Ils déterminent ainsi à quelles informations les gens ont accès et, par conséquent, tentent de configurer significativement – sans y parvenir complètement – les pratiques sociales.

Le rôle sociotechnique des logiciels selon les théories de la conception

Le système de valeurs reproduit dans les logiciels et mis en œuvre au moment de l'interaction avec les usagères et usagers s'inscrit à l'étape de la conception des systèmes algorithmiques. Les choix effectués lors de la création de ces logiciels renvoient à ce que Gillespie (2014) nomme les modèles d'inclusion. Les éléments exclus ou inclus dans ces technologies (données, structures, techniques, etc.) mettent en évidence l'intervention sémantique et politique des algorithmes. Ces calculs décident du classement des informations et de la façon dont les données seront traitées. Une analyse des forces sociales et culturelles contribuant à la conception de ces systèmes (Manovich, 2013) permet la compréhension de leur dimension tant sociale que technique. Les théories de la conception, dont la théorie de l'acteur-réseau (Akrich, 2006a), permettent d'envisager la création de ces systèmes sociotechniques, telle une construction effectuée par et dans un réseau d'acteurs. Les intérêts de ces acteurs, qu'ils relèvent de la conception, des usages ou des environnements, s'insèrent dans une opération d'opposition et de rapprochements où des négociations effectuées entre eux stabilisent le système sociotechnique (Akrich, 2006a; 2006b). Dans cette opération, les usages imaginés sont configurés dès la conception du système technique (Woolgar, 1990).

Issue des théories de la conception, la notion de « configuration des usagers » (Woolgar, 1990) effectuée par et dans les objets techniques peut être associée à la notion d'opération d'inscription présente dans la conception des technologies. L'opération d'inscription fait référence à l'anticipation des conditions d'utilisation des objets techniques effectuée dès la conception (Jauréguiberry et Proulx, 2011). Les règles de fonctionnement adoptées correspondent à ce qu'Akrich (1998) appelle les scripts des entités techniques composant l'objet. Aussi bien la représentation de l'utilisateur

imaginé que les usages prescrits sont inscrits dans ces scripts par les personnes chargées de la conception. Ces scripts rendent possible l'action et guident l'usage des objets et des interfaces. Ils peuvent donc être envisagés comme des objets d'analyse puisqu'ils dévoilent des indices de la stratégie employée au moment de la conception (Bardini, 1996). Les usages envisagés ou prescrits sont également encouragés par des plans d'action qui tentent de les coordonner (Akrich, 2006a). Ces plans donnent des orientations explicites concernant l'usage de l'objet technique, tels que les guides d'usage et les sites Web associés à l'objet.

Un cadre d'analyse adapté à l'étude des logiciels

L'étude des scripts (Akrich, 2006a) permet notamment d'entamer un processus de repérage de ce que nous avons défini comme étant la configuration de l'usage (Woolgar, 1990), sans passer nécessairement par une interaction directe avec les personnes travaillant à la conception du logiciel à l'étude ou avec celles qui en font usage. Néanmoins, à bien des égards, les logiciels se présentent de manière opaque. La compréhension des logiques prescriptives établies implique de prêter attention à au moins trois de leurs dimensions : la « boîte noire » les constituant ; les signes composant l'interface offerte aux usagers et usagères ; et le plan de coordination des usages appliqué par les acteurs de sa conception.

Un tel cadrage analytique reposant sur un ensemble mixte de méthodes compréhensives de recherche permet de leur redonner une intelligibilité. Dans un premier temps, il apparaît nécessaire d'offrir une description détaillée du processus de construction des composantes de ces systèmes algorithmiques, des relations créées entre elles dans cet environnement et des résultats produits par ces relations pour comprendre la logique prescriptive de la « boîte noire ». Dans un deuxième temps, la manifestation de cette logique sur l'interface de l'objet technique permet de repérer les usages encouragés dans les scripts, notamment par l'analyse des formes sémiotiques constituant l'interface logicielle. Enfin, l'analyse des plans de coordination des usages permet de reconstruire les prescriptions d'usage imaginées dans les instructions explicites et leurs justifications.

L'étude des algorithmes ou la « boîte noire »

Dans le cas de la recherche menée sur Netflix, deux corpus ont été analysés dans l'étude des algorithmes, les « boîtes noires » de ce service. Le premier consiste en un texte décrivant la solution algorithmique retenue par Netflix lors du concours Netflix Prize et appliquée dans son système de recommandation. En 2006, l'entreprise a offert le prix d'un million de dollars à qui pourrait améliorer de 10 % l'efficacité de Cinematch, son système de recommandation à l'époque. Le deuxième corpus a été composé d'une sélection de publications du site Web *The Netflix Tech Blog*, un site créé par l'entreprise en 2010, visant à présenter les perspectives, les décisions et les défis de Netflix dans la gestion de son logiciel. Ces deux corpus ont permis de retracer la trajectoire des algorithmes depuis la fin du concours Netflix Prize, en 2009, jusqu'en 2015, l'année précédant la réalisation de notre étude. À cette étape, des comptes rendus chronologiques et des tableaux ont contribué à l'analyse du fil des événements.

Dans le cas du système de recommandation de Netflix, les composantes (profils des usagers et contenus) sont associées à partir d'un système de valeurs (modélisation) afin de générer des listes de recommandations (résultats) pour chaque profil existant dans le compte. Les profils créés par les gens correspondent à leur représentation inscrite dans le logiciel. Les autres composantes, comme les contenus du catalogue de Netflix, sont associées à cette représentation de la personne dans un processus de modélisation. Ce dernier génère, à son tour, les résultats ou recommandations.

L'étude des textes du concours Netflix Prize et du *Tech Blog* nous a permis de construire une grille d'analyse centrée sur le repérage du processus d'attribution des caractéristiques aux composantes du système et des règles existantes dans le processus de modélisation. L'analyse des mécanismes de profilage de la « boîte noire » repose sur la compréhension de la conception des composantes, de la définition de leur identité et de l'application des règles établies. Elle permet de définir les contours des composantes et de l'environnement inscrit (Akrich, 2006a) dans l'objet technique. La compréhension de la logique prescriptive des systèmes algorithmiques peut être ainsi examinée à partir de quatre catégories d'analyse, issues de ce processus de conception : la représentation de l'individu qui fera usage de l'objet technique (attribution explicite de caractéristiques à cet usager imaginé), le classement des autres compo-

santes (attribution de caractéristiques aux éléments avec lesquels cette représentation sera mise en relation), la modélisation des résultats (système de valeurs et de poids) et les résultats.

Compte tenu des transformations régulièrement opérées dans les logiciels – dont les algorithmes –, comme celles apportées au système PageRank de Google³ au fil du temps (Cardon, 2013), on remarque que des changements peuvent se produire dans l'environnement inscrit dans la « boîte noire ». La trajectoire des modifications de ces systèmes se présente comme un élément pertinent dans l'analyse de l'algorithme. En effet, l'intervention des acteurs, dans le cadre, notamment, du changement de stratégie de conception et de l'action d'utilisateurs « bricoleurs » (Cardon, 2013), influe sur ces transformations. Repérer la trajectoire des transformations de la « boîte noire » représenterait un moyen de mieux comprendre l'utilisateur configuré comme une version de l'« utilisateur effectif » (Jauréguiberry et Proulx, 2011) produite par un système constamment modifié. En même temps, la configuration de l'utilisateur serait comprise comme un processus qui se construit tout au long du cycle de vie d'une technologie.

Les précautions pour l'analyse d'une « boîte noire »

L'analyse des stratégies de configuration des usages mises en place dans les « boîtes noires » comporte des défis particuliers en ce qui concerne le langage algorithmique, compte tenu de sa complexité technique. Une collaboration interdisciplinaire (spécialistes en informatique, en intelligence artificielle, etc.) s'avère nécessaire dans ce type de recherche. Des lectures exploratoires des corpus permettent de bien saisir les expertises nécessaires dans la recherche, telles que l'intelligence artificielle et le génie logiciel. Dans le cas de la recherche sur Netflix, nous avons contacté un professeur et une doctorante impliqués dans ces domaines afin de leur présenter un résumé de la recherche et le type de collaboration qui serait établi.

3. Cardon (2013) s'est particulièrement intéressé à la fragilisation de l'ancien algorithme de Google, PageRank. L'action d'internautes qui pouvaient, par exemple, modifier l'offre de résultats du moteur de recherche de la plateforme, a abouti à la restructuration du système et à la création de l'algorithme Panda. Les comportements rusés étaient détectés et punis par l'entreprise. Panda a été ainsi créé comme une réaction de l'entreprise aux actions des utilisateurs « bricoleurs » et aux limites du PageRank initial.

D'ailleurs, des connaissances de base en informatique ou en *machine learning* (apprentissage automatique) représentent des éléments indispensables pour qui désire étudier les algorithmes. Des ouvrages et d'autres types de contenus vulgarisant le sujet ainsi que des rencontres avec des spécialistes du domaine de l'informatique représentent une porte d'entrée pour la compréhension des mécanismes de la « boîte noire ». Lors du décodage des matériaux, l'assistance de ces spécialistes s'avère indispensable puisqu'ils jouent un rôle d'interprètes⁴ à cette étape. La création d'un lexique pour les concepts techniques peut contribuer de manière significative aux analyses, car ces concepts peuvent revenir au cours de la recherche. La révision des résultats par l'équipe de recherche demeure essentielle pour éviter les mécompréhensions et valider les résultats de recherche. De plus, l'opacité des systèmes algorithmiques (Burrell, 2016) impose des limites aux recherches visant à les analyser, puisque les mesures de sécurité des environnements en ligne des entreprises et la protection des secrets de ces entreprises peuvent restreindre l'accès à la « boîte noire ».

L'étude des signes sur l'interface des logiciels

L'analyse sémio-communicationnelle (Bonaccorsi, 2013; Candel *et al.*, 2012) de l'interface du logiciel⁵ représente un outil permettant de comprendre la configuration de l'usage proposée par les responsables de sa conception. Cette configuration se manifeste à travers le code symbolique inscrit dans l'interface de l'objet technique (Bardini, 1996). La personne usagère, l'objet et les concepteurs établissent un rapport concret grâce à ce code. L'analyse du script de l'interface d'un logiciel permet ainsi de comprendre cette médiation physique et symbolique.

L'analyse sémio-communicationnelle de l'interface du logiciel se concentre sur la découverte, la catégorisation et la compréhension des signes visibles manifestés dans les contenus textuels et visuels de cet objet. Ces signes composent la grammaire éditoriale (Candel *et al.*, 2012) des objets techniques. Ils permettent l'action et stimulent les usages

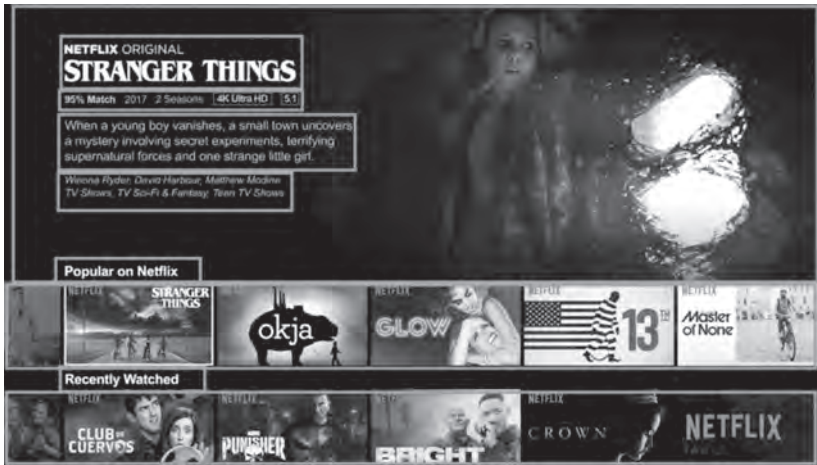
4. Nous tenons à remercier la collaboration du professeur Mounir Boukadoum du Département d'informatique de la Faculté des sciences de l'UQAM et de la doctorante en intelligence artificielle à l'UQAM, Ange Tato, dans le cadre de la recherche sur Netflix.

5. Nous avons analysé la version pour les ordinateurs qui est accessible par les navigateurs. Les autres versions de Netflix, comme celle pour les appareils mobiles par exemple, n'ont pas fait l'objet de cette étude.

(Bonaccorsi, 2013). Ces formes (Candel *et al.*, 2012) exercent une fonction à la fois pragmatique et normative (Grignon, 2015). Les signes exercent également une fonction symbolique d'ordre connotatif, puisqu'ils correspondent à une sorte d'identifiant qui concrétise un plan d'anticipation des usages (Piment, 2015).

FIGURE 16.1

Illustration de la page d'accueil de Netflix (2018) indiquant quelques éléments textuels et des images trouvés dans l'interface



Source: Site Presse de Netflix

Le classement de ces signes visibles sur l'interface peut être formalisé sous la forme d'une grille précisant la description de chaque catégorie (voir tableau 16.1). Le « codage ou identification » des signes représente un premier processus d'étiquetage et de désignation des éléments analysés (ligne 1). Compte tenu de la dynamique des innovations (Candel *et al.*, 2012), le fait d'indiquer la date du moment où les captures d'écran ont été faites représente un moyen de saisir les modifications des signes au fil du temps (ligne 2). La catégorisation des formes selon leur contenu ou leur fonction dans l'interface (ligne 3) s'avère nécessaire lors de l'établissement des objectifs fixés à la conception (Piment, 2015) qui se rendent visibles à travers les signes. Dans le cas de Netflix, par exemple, les rangées où les contenus sont disposés peuvent être catégorisées avec l'étiquette « recommander les contenus » à la ligne 3 du tableau.

La description détaillée de l'apparence des signes peut être réalisée sur trois plans : textuel (les textes composant l'interface, comme les titres des rangées de contenus, dans le cas de Netflix), imagé (les images composant l'interface, comme les vignettes reproduisant les affiches des films) et iconographique (les icônes composant l'interface et permettent les actions, comme le bouton «jouer») (lignes 4, 5 et 6). Une description détaillée des positions des signes (Grignon, 2015), comme celle du bouton «jouer» sur l'interface, peut préciser l'emplacement des éléments indiqués aux lignes 7, 8 et 9. Le cadre de l'usage est couvert par les lignes où est indiquée la désignation des actions possibles, des possibilités d'engagement (intervention de l'utilisateur sur l'objet) et des suggestions d'action (Piment, 2015) (lignes 11, 12 et 13).

Des textes narratifs, comme des fiches de synthèse, des tableaux et des images, contribuent à mieux structurer et à présenter la logique prescriptive. Les fiches de synthèse contribuent particulièrement à établir la liste d'éléments saillants issus de la grille d'analyse, ce qui rend plus facile la formalisation des résultats de recherche. Les tableaux exposant ces

TABLEAU 16.1

Grille d'analyse sémio-communicationnelle (interface du logiciel)

N	Catégories
1	Codage /identification
2	Date de la capture d'écran
3	Catégorie/fonction
4	Éléments textuels
5	Images
6	Iconographie (icônes et formes)
7	Disposition éléments textuels (espace)
8	Disposition images
9	Disposition icônes (espace)
10	Représentation de l'utilisateur
11	Actions possibles
12	Possibilités d'engagement avec l'objet
13	Suggestions d'actions

éléments permettent de les hiérarchiser tandis que les images illustrent les informations contenues dans les tableaux. Ces outils peuvent aussi servir à la révision et à la présentation des résultats de recherche.

L'étude du plan de coordination des actions

Étant donné que l'objet technique peut être utilisé de différentes manières, des stratégies de coordination des usages sont déployées afin de favoriser le bon fonctionnement de l'objet et d'éviter les usages non désirés (Akrich, 2006a). Pour ce faire, l'univers des usagers et usagères ainsi que les éléments pouvant représenter des obstacles à l'accomplissement de l'action sont imaginés lors de la conception. Ces problèmes sont anticipés et utilisés dans l'élaboration des matériels ou des outils associés à l'objet technique, tels que les guides d'usage et les sites Internet, où les bons usages sont expliqués et suggérés. Ces stratégies de coordination des actions contribuent à une modélisation ou une sorte de « préconfiguration » des usages. Elles visent à réduire au minimum les initiatives des gens et à leur enseigner un programme d'action (Akrich, 2006a). Ces stratégies se manifestent par un code symbolique présent dans les entités techniques composant le plan de coordination des usages, ce qui relève encore d'un cadre sémiotique à analyser dans cette troisième étape.

La reprise de la grille d'analyse sémio-communicationnelle, déjà utilisée dans l'étude des signes de l'interface de Netflix, a permis de repérer les dimensions physique et symbolique du script du site Web de ce service. Toutefois, quelques ajustements ont été nécessaires. Les lignes concernant l'emplacement des éléments textuels, des images et des icônes (lignes 6, 7 et 8) ont été supprimées, car ce type de script ne compose pas directement l'objet technique lui-même. Dans le cas de Netflix, le site Web de l'entreprise représente la porte d'entrée du service en permettant la création des comptes et l'accès à la plateforme. Afin de catégoriser les instructions données et leur objectif en rapport avec la configuration de l'usage (assistance à la clientèle, suggestions d'usage, etc.), la ligne « nature du contenu informationnel » (Piment, 2015) a été ajoutée à la grille (ligne 3).

À nouveau, des textes narratifs, comme des fiches de synthèse, des tableaux et des images peuvent contribuer à la structuration et à la présentation des résultats de recherche.

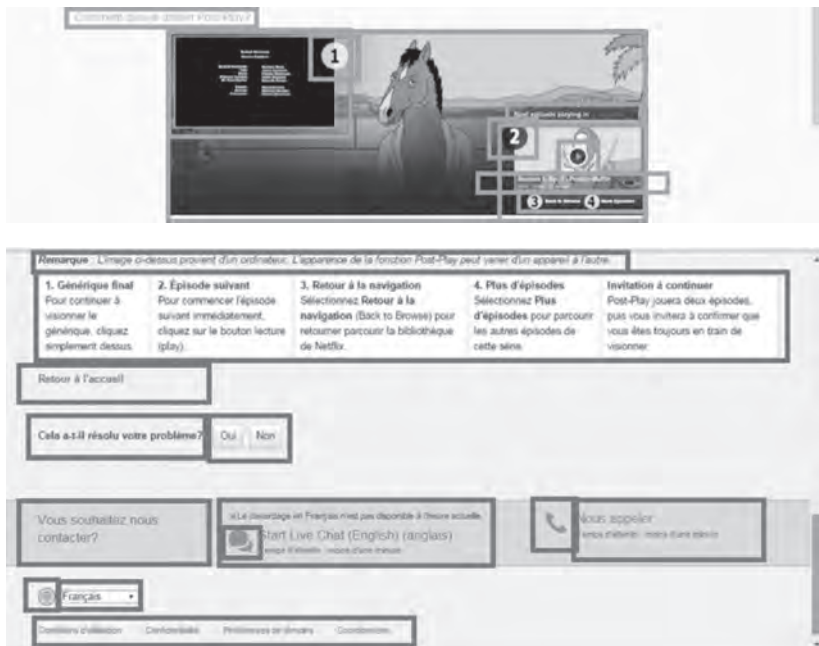
TABLEAU 16.2

**Grille d'analyse sémio-communicationnelle
(plan de coordination des actions)**

N	Catégories
1	Codage /identification
2	Date de la capture d'écran
3	Nature du contenu informationnel
4	Éléments textuels
5	Images
6	Iconographie (icônes et formes)
7	Représentation de l'utilisateur
8	Actions possibles
9	Possibilités d'engagement avec l'objet
10	Suggestions des actions

FIGURES 16.2 ET 16.3

**Captures d'écran du Centre d'aide du site Web de Netflix (02/03/2016)
indiquant quelques éléments textuels, des images et des éléments
iconographiques**



La version configurée des usages de Netflix

La configuration des usages par Netflix découle de quatre éléments principaux, à savoir la personnalisation, la création des catégories de contenu, l'adaptation et les rétroactions. Dans le *Tech Blog*, l'interface du logiciel et le site Web, la personnalisation est présentée comme étant un aspect élémentaire du fonctionnement du logiciel. Dès la connexion au compte, elle peut être reconnue dans les icônes et les éléments textuels liés aux profils des personnes utilisant le service. En suggérant de réserver un profil spécifique à chaque personne du foyer, par exemple, Netflix enseigne une façon de consommer les contenus et met en relief son modèle d'affaires. En expliquant des aspects de la catégorisation des contenus, l'entreprise met en lumière ses efforts pour personnaliser son service. Dans les titres des rangées de contenus, comme dans les rangées «*Because you watched* [titre d'un contenu visionné] », par exemple, Netflix explique les listes présentées à l'écran et, par conséquent, expose les travaux de son système de recommandation. Pour l'entreprise, ce travail d'argumentation sert à favoriser la confiance des gens dans son système. L'usage configuré de Netflix est ainsi plus ou moins conscient du traitement des données et de la création des recommandations.

Selon les textes du *Tech Blog*, l'adaptation du logiciel à l'environnement dans lequel les usagers se retrouvent permettrait à Netflix d'avoir une version configurée plus proche de la réalité. Les données de géolocalisation, la qualité de la bande passante et le choix de la langue contribueraient à une meilleure précision de la version des usagères et usagers produite par le système. Pour l'entreprise, l'adaptation du logiciel aux différents appareils favorise une meilleure expérience de son service et évite que des erreurs de configuration de l'interface, comme l'affichage de listes incomplètes de recommandations, nuisent aux travaux du système de recommandation. En approuvant ou en refusant leur version configurée, les usagères et usagers fournissent des exemples positifs (ce qui fonctionne) et négatifs (ce qui ne fonctionne pas) à Netflix.

Outre l'historique des usages, les données déclarées par les personnes se servant des outils de rétroaction contribuent activement à la création des recommandations. La rangée *My list* et le système de notation de contenus en représentent quelques exemples. Ces outils permettraient de mieux saisir les vrais intérêts des gens et ainsi de faire des recommandations encore plus pertinentes. La fonction *Post-play*, qui permet d'enchaîner

automatiquement les contenus, génère un volume important de rétroactions lors des marathons de visionnement. Ce mode de consommation fournirait aussi des indices clairs à propos des préférences des gens. L'écoute en rafale est ainsi en accord avec une logique de récolte de données massives, surtout de données précisant des informations sur l'« usager effectif ».

À partir des travaux de son logiciel et de la mise en discours des prescriptions des usages, Netflix tente d'amener les usagers à se conformer à sa logique de fonctionnement. Plus on regarde de contenus sur Netflix, plus de données sont fournies au système, et plus l'offre de recommandations est pertinente. En adoptant les pratiques de visionnement prescrites par l'entreprise, les usagers auraient une expérience chaque jour plus exceptionnelle de son service.

L'usage configuré inscrit dans les algorithmes relève de la complexité de l'objet technique lui-même ainsi que des rapports établis entre les usagers et usagères, les objets techniques, les personnes chargées de sa conception et l'environnement dans lequel ils se retrouvent (Akrich, 2006a). En mobilisant les théories de la conception et des travaux provenant des *software studies*, un cadrage analytique a été proposé pour explorer la configuration des usages effectuée par les logiciels. Afin d'étudier la représentation des individus (Bardini, 1996) et les usages prescrits (Jauréguiberry et Proulx, 2011) inscrits dans l'objet technique, des analyses des scripts des objets techniques sont effectivement nécessaires. Nous avons proposé ici de mener celles-ci en trois temps, en analysant à la fois la « boîte noire », l'interface et les actions suggérées sur le site Web. Selon les résultats que nous avons obtenus dans notre étude de cas sur Netflix, le cumul de ces trois degrés d'analyse permet de mettre à jour la logique prescriptive de la « boîte noire » et sa manifestation dans l'interface ainsi que les instructions explicites données par les concepteurs.

Les algorithmes décident de ce à quoi les gens ont accès, et les critères utilisés demeurent obscurs et peu accessibles (Beer, 2013). Ces technologies tentent également d'anticiper les choix des gens en leur créant une « version configurée ». Les calculs algorithmiques calibrent les représentations des personnes faisant usage des technologies ainsi que les usages prescrits afin d'aligner l'« usager effectif » et sa version configurée. Les effets des algorithmes sur la vie de tous les jours et leur rôle dans la culture contemporaine représentent des sujets encore à explorer. D'ailleurs, l'analyse des

logiques prescriptives appliquées par et dans ces technologies contribue à la compréhension de leur évolution sociotechnique ainsi que des enjeux sociaux, politiques et économiques découlant de la configuration des usages. Des recherches utilisant un appareillage méthodologique centré sur leur conception permettent de cerner les dimensions prescriptives de cas particuliers, comme des études sur les algorithmes d'autres plateformes comme Facebook et Google, entre autres. Cet appareillage fournit aussi des outils qui permettent de comprendre des contextes plus larges, comme celui dans lequel la relation dialogique entre les stratégies de gestion de l'information et les transformations des pratiques se construit.

Références

- AKRICH, Madeleine, « La construction d'un système sociotechnique. Esquisse pour une anthropologie des techniques », dans Madeleine AKRICH, Michel CALLON et Bruno LATOUR (dir.), *Sociologie de la traduction: textes fondateurs*, Presses des Mines, 2006a, p. 109-134.
- AKRICH, Madeleine, « Les objets techniques et leurs utilisateurs. De la conception à l'action », dans Madeleine AKRICH, Michel CALLON et Bruno LATOUR (dir.), *Sociologie de la traduction: textes fondateurs*, Presses de Mines, 2006b, p. 179-199.
- AKRICH, Madeleine, « Les utilisateurs, acteurs de l'innovation », *Éducation permanente*, vol. 134, 1998, p. 79-90.
- BARDINI, Thierry, « Changement et réseaux socio-techniques: De l'inscription à l'affordance », *Réseaux*, vol. 14, n° 76, 1996, p. 125-155.
- BEER, David, « Algorithms: Shaping tastes and manipulating the circulations of popular culture », dans *Popular Culture and New Media: The Politics of Circulation*, Palgrave Macmillan, 2013, p. 63-100.
- BONACCORSI, Julia, « Approches sémiologiques du Web », dans Christine BARATS (dir.), *Manuel d'analyse du Web en sciences humaines et sociales*, Armand Colin, 2013, p. 125-146.
- BURRELL, Jenna, « How the machine "thinks": Understanding opacity in machine learning algorithms », *Big Data & Society*, vol. 3, n° 1, 2016, p. 1-12.
- CANDEL, Étienne, Valérie JEANNE-PERRIER et Emmanuël SOUCHIER, « Petites formes, grands desseins. D'une grammaire des énoncés éditoriaux à la standardisation des écritures », dans Jean DAVALLON (dir.), *L'économie des écritures sur le Web*, vol. 1: *traces d'usage dans un corpus de sites de tourisme*, Hermès-Lavoisier, 2012, p. 165-201.
- CARDON, Dominique, *À quoi rêvent les algorithmes. Nos vies à l'heure des big data*, Éditions du Seuil, 2015.
- CARDON, Dominique, « Dans l'esprit du PageRank. Une enquête sur l'algorithme de Google », *Réseaux*, vol. 177, n° 1, 2013, p. 63-95.

- DRUMOND, Gabrielle Silva MOTA, *La configuration des usages sur Netflix: le système de recommandation Cinematch et la représentation de l'utilisateur*, mémoire de maîtrise en communication, Université du Québec à Montréal, 2016.
- GILLESPIE, Tarleton, « The relevance of algorithms », dans Tarleton GILLESPIE, Pablo L. BOCZKOWSKI et Kirsten A. FOOT (dir.), *Media Technologies: Essays on Communication, Materiality, and Society*, The MIT Press, 2014, p. 167-193.
- GRIGNON, Thomas, « L'expertise communicationnelle au prisme de ses instruments. L'exemple de Google Analytics », *Les Cahiers du RESIPROC*, n° 3, 2015, p. 23-47.
- HALLINAN, Blake et Ted STRIPHAS, « Recommended for you: The Netflix Prize and the production of algorithmic culture », *New Media & Society*, vol. 18, n° 1, 2016, p. 117-137.
- JAURÉGUIBERRY, Francis et Serge PROULX, *Usages et enjeux des technologies de communication*, Érès, 2011.
- JENNER, Mareike, « Is this TVIV? On Netflix, TVIII and binge-watching », *New Media & Society*, vol. 18, n° 2, 2016, p. 257-273.
- KITCHIN, Rob et Martin DODGE, *Code/space: Software and Everyday Life*, The MIT Press, 2011.
- LAROUSSE, « Algorithme », *Dictionnaire*, [En ligne], [<https://www.larousse.fr/dictionnaires/francais/algorithme/2238?q=algorithme>], consulté 8 avril 2018.
- MANOVICH, Lev, *Software Takes Command*, Bloomsbury Publishing, 2013.
- NETFLIX TECHNOLOGY BLOG, « Netflix recommendations: Beyond the 5 stars (Part 2) », *Medium*, 20 juin 2012, 2 avril 2018.
- PIMENT, Hélène, « Modèle communicationnel d'un réseau socionumérique d'entreprise », *Les Cahiers du RESIPROC*, n° 3, 2015, p. 100-123.
- SHARMA, Amit, « Third-Party recommendation systems industry: Current trends and future directions », *SSRN Scholarly Paper*, Social Science Research Network, Rochester, 2013.
- STRIPHAS, Ted, « Algorithmic culture », *European Journal of Cultural Studies*, vol. 18, n° 4-5, 2015, p. 395-412.
- WOOLGAR, Steve, « Configuring the user: The case of usability trials », *The Sociological Review*, vol. 38, n° S1, 1990, p. 58-99.

POSTFACE

Une nécessaire politique des méthodes pour la transition numérique

Serge Proulx

En rendant compte des diverses facettes des approches méthodologiques pour la recherche en sciences de l'information et de la communication en temps de transition numérique, cet ouvrage témoigne d'un réel dynamisme du domaine d'étude et plusieurs de ses propositions novatrices méritent d'être prises en considération. Son fil conducteur se trouve dans la pertinence des approches qualitatives pour étudier aujourd'hui les activités en ligne. Il s'agit d'un postulat épistémologique fort à l'heure des illusions tracées à l'horizon des données massives (*big data*) – la prétention de vouloir faire l'économie des théories en sciences sociales n'en étant pas la moindre. Ce mot à la mode est présent aussi bien dans le discours des médias que dans les considérations de certaines communautés de recherche axées sur les sciences de l'information et de la communication. Par contraste, et en complémentarité, il faut souligner les transformations significatives, et profondes, du domaine des approches quantitatives à l'ère du numérique.

Je voudrais insister sur quelques thématiques transversales pouvant aider à penser le champ des méthodes dans le contexte contemporain de la transition de nos sociétés vers le tout numérique. Tout se passe comme si la très grande majorité des pratiques d'information et de communication se trouvaient automatiquement associées à un univers technique numérique (nature des supports dits intelligents; encodage algorithmique des signaux et des informations; capture, accumulation et monétisation des

traces numériques des trajectoires des sujets communicants). Les pratiques de communication interhumaine non médiatisée par la technique numérique sont une denrée rare dans les sociétés contemporaines. Aujourd'hui, les plateformes numériques et les machines informationnelles deviennent le passage obligé de toutes les transactions interhumaines. Cet état des choses suppose une adaptation scientifique des méthodes, voire une rupture épistémologique – il y a là matière à débat, comme Boullier le laisse entendre dans le présent ouvrage – dans le *design* des dispositifs méthodologiques utilisés en sciences sociales pour rendre compte de ces pratiques numériques d'information et de communication.

Le terme « méthode » ne désigne pas exclusivement le domaine des techniques de collecte et d'analyse des données. Il renvoie encore et surtout au champ des méthodologies, c'est-à-dire à l'articulation raisonnée et pragmatique d'un choix spécifique de méthodes qui impliquent des présupposés épistémologiques, et qui induisent en même temps des conséquences politiques dans l'environnement économique et social où elles se déploient (voir le chapitre de Millerand, Myles et Proulx dans le présent ouvrage). Aussi, à un moment de transition généralisée au tout numérique, il est justifié d'évoquer la nécessité de repenser une « politique des méthodes » définie comme l'intelligence stratégique concernée par le pouvoir social de la connaissance obtenue par les recherches en sciences humaines et sociales, qui ont comme objet les activités privées et publiques dans la sphère sociale. Or il est d'abord important de noter que ces recherches utilisent des méthodologies spécifiques qui ont des conséquences politiques.

Un regard politique et épistémique sur les choix méthodologiques des recherches met en évidence le fait que la connaissance produite – qui se présente sous la forme de données – n'est pas politiquement neutre. Elle s'avère un construit épistémique encastré dans des choix apparemment « techniques » correspondant à certains « besoins » particuliers des organisations et des infrastructures de production. Ainsi, certaines techniques et méthodes sont davantage en résonance avec les écosystèmes informationnels où ces recherches se réalisent. Nos sociétés contemporaines sont en effet organisées selon des règles économiques propres. En l'occurrence, celles du capitalisme informationnel et cognitif, c'est-à-dire d'un mode de production dont la valorisation économique est fondée précisément sur l'accumulation et la monétisation de données transac-

tionnelles obtenues dans le cadre d'échanges et d'activités en ligne qui transitent par les plateformes numériques. L'on peut s'attendre à ce que les choix qui tendent à la production de données « quantifiées » soient valorisés *a priori* dans un tel système économique. La définition d'une politique des méthodes dans un tel contexte rejoint, dans sa finalité, la mise en place d'une imagination méthodologique radicalement novatrice qui s'articule autour de la critique de la numérisation généralisée des sociétés contemporaines.

La recomposition du champ des méthodes en sciences sociales

Dans un texte de 2007 préfigurant ce qu'on appellera plus tard les données massives, Savage et Burrows (2007) diagnostiquaient l'avènement d'une crise de reconnaissance, voire de légitimité affectant la sociologie empirique anglo-saxonne, dépassée par l'émergence d'un nouveau type de données sociales produites par l'industrie privée à l'insu des sciences sociales universitaires. Or, douze années plus tard, force est de constater que ce diagnostic s'est avéré fondé. En effet, tout se passe aujourd'hui comme si les sciences sociales universitaires n'étaient plus le « point de passage obligé », comme l'évoquait Latour (2007a), dans la production exclusive de données pertinentes témoignant du fonctionnement de la sphère sociale. Les sciences sociales universitaires auraient perdu le monopole du contrôle et de la diffusion des méthodes empiriques de recherche, au profit de nouveaux acteurs industriels, notamment parce que ces agents économiques produisent eux-mêmes et de manière autonome les infrastructures logicielles auxquelles ils recourent dans leur activité de valorisation économique. Les nouvelles entreprises numériques n'auraient plus besoin, pour leur fonctionnement, d'utiliser les données sociales produites par la sociologie traditionnelle pour connaître les caractéristiques de leurs clientèles.

Le capitalisme informationnel et cognitif (Proulx, 2016) – c'est-à-dire une forme d'organisation économique apparue dans le dernier quart du *xx^e* siècle et dans laquelle la production de la valeur économique est fondée sur la circulation et l'accumulation de l'information et de la connaissance – induit l'usage et le déploiement d'un nouveau type de données, à savoir les données sociales transactionnelles. Or, ces données sont produites directement par les entreprises publiques, parapubliques et privées, indus-

rielles et commerciales, qui gèrent une grande partie des activités qui occupent le temps et mobilisent aujourd'hui l'attention des personnes (travail, consommation, communication, information, loisirs, jeux, divertissement). Ces données transactionnelles sont générées à partir des listes d'envoi (*mailing lists*) aux clients, mais aussi des traces de chacune de leurs transactions (demande d'information, achat, service après-vente), de leur usage des cartes de fidélité (permettant de constituer des corrélations entre les achats d'une même personne), des listes d'abonnement à des services, etc. Les caractéristiques des groupes semblables peuvent ainsi être corrélées pour produire des scénarios prédictifs d'achats futurs de la part de segments de clientèles et par conséquent, des stratégies et des tactiques publicitaires précisément ciblées. Ainsi, les analyses approfondies de ces corpus de données transactionnelles ont révélé un résultat surprenant (non intuitif), à savoir l'importance décisive de la variable « localisation géographique » des transactions pour permettre la construction des scénarios prédictifs des comportements futurs des personnes (Savage et Burrows, 2007, p. 892).

Dans une organisation sociale où les infrastructures de production et de consommation reposent de plus en plus sur les technologies de l'information et de la communication, ces données transactionnelles sont produites de manière automatique. Ce sont les traces informationnelles laissées automatiquement sur les plateformes numériques par les personnes chaque fois que ces dernières effectuent des transactions avec les entreprises connectées. Il faut comprendre que la notion de « transaction » est ici une catégorie cognitive aux contours définitionnels très larges puisque le simple fait pour une personne d'exister dans un environnement informationnel laisse des traces, qu'elle le veuille ou non (Latour, 2007b). L'on pourrait parler ici d'un écosystème informationnel agissant sous l'emprise des technologies de surveillance au quotidien. Notons aussi que cette surveillance s'accroît du fait même de la mise en place dans l'environnement de multiples objets connectés (« Internet des objets ») qui captent les traces des multiples transactions des individus avec ces objets et les renvoient vers des infrastructures logicielles centralisées (bases de données).

Avant l'arrivée d'Internet et des techniques numériques de captation, de transmission et d'accumulation de traces transactionnelles, les sciences sociales universitaires nous avaient habitués à travailler avec d'autres types

de données quantitatives et qualitatives. D'abord, et surtout, nous connaissons les données quantitatives d'enquêtes par questionnaires administrés à l'aide de sondages auprès d'échantillons représentatifs de populations ciblées. Ensuite, les données produites par le recours aux techniques qualitatives d'entretien (dirigé; semi-directif; non directif, appelé aussi compréhensif ou en profondeur; histoires et récits de vie; récits de pratiques; entrevues de groupe et groupes de discussion [*focus groups*] et aux techniques d'observation des activités humaines en contexte (observation orientée ou flottante, participante ou non). Mentionnons aussi les enquêtes ethnographiques proprement dites, les analyses de réseaux sociaux non numériques (*social network analysis*) et les analyses de cas (*intensive case studies*) qui concourent à la construction de corpus importants de matériaux denses, qualitatifs et quantitatifs, pensés en contexte (*thick description*), sans oublier les données obtenues par les diverses analyses documentaires (examen d'archives, statistiques sociodémographiques, analyse de discours, comparaisons historiques, etc.).

Les méthodes d'enquête par sondage ont constitué une innovation importante issue des sciences sociales, en particulier à partir des décennies 1940-1950. En se couplant aux statistiques inférentielles et à la théorie des probabilités, les méthodes d'enquête par sondage ont acquis progressivement une robustesse scientifique. À partir du moment où un échantillon aléatoire est bien construit et bien calibré statistiquement, il peut prétendre représenter la totalité de la population visée par l'enquête ainsi que les divers segments sociodémographiques qui la constituent. Il se trouve ici une importante économie d'échelle du fait qu'il suffit de n'observer qu'un échantillon de la population pour avoir la possibilité de généraliser les résultats à l'ensemble de la population. Ces approches statistiques se sont progressivement raffinées au moyen de techniques d'estimation des caractéristiques des non-répondants. Jusqu'aux années 1990, les enquêtes par questionnaires administrés par sondages téléphoniques ou en personne demeuraient une méthode quantitative privilégiée par les sciences sociales universitaires.

À partir de cette période toutefois (la décennie 1990), alors même que les études de marché avaient connu un essor important, les taux de réponses aux sondages téléphoniques se sont mis à chuter, les personnes trop sollicitées se sentant davantage concernées par la préservation de leur intimité. Depuis, des techniques hybrides ont été mises au point, en

calibrant les sous-échantillons selon le support mobilisé pour joindre la personne (téléphone fixe, téléphone mobile, accès par le Web). Idéalement, et en résonance avec le principe de la triangulation des méthodes (explicité notamment dans le chapitre de Roginsky dans le présent ouvrage), les dispositifs d'enquête les plus performants combinent ces techniques quantitatives avec des techniques qualitatives d'entretien (individuels et de groupe) ou d'observation destinées à produire des données complémentaires fournissant un matériel pertinent menant à une compréhension approfondie des significations sociales que les sujets attribuent à leurs actions lorsqu'ils laissent libre cours à une réflexivité concernant leurs propres perceptions et comportements (voir les chapitres de Péliissier sur l'autoconfrontation et de Gallant, Pastinelli, Latzko-Toth et Labrecque sur la visite commentée des historiques personnels). En d'autres termes, il s'agit d'une explication de leurs propres valeurs et de leur vision du monde.

L'avènement d'Internet et la généralisation de l'usage des techniques numériques de captation automatique des traces de transactions entre les humains et les machines informationnelles changent l'état des choses. La démultiplication des données transactionnelles produites par les plateformes numériques rend relativement obsolètes les données quantitatives produites au moyen des enquêtes par sondages. Il s'avère en effet que ces corpus de données transactionnelles recouvrent la totalité des activités qui se sont déroulées, et non seulement un échantillon de ces activités. Certains spécialistes en sciences des données (*data science*) prétendent ainsi pouvoir faire l'économie d'un recours aux techniques de la représentativité statistique puisqu'il s'agit de travailler directement sur les traces d'activités de populations entières. Ces spécialistes mettent également en avant le fait qu'ils évitent de recourir à un matériel « déclaratif », c'est-à-dire construit selon les déclarations subjectives des personnes approchées par le moyen des questionnaires ou des entretiens. Une partie des spécialistes en sciences des données vont même jusqu'à prétendre que leurs techniques numériques leur permettraient d'avoir accès à la « réalité objective » des activités des agents en évitant ainsi la médiation des perceptions subjectives, ce qui constitue selon nous une illusion positiviste grossière (Proulx et Rueff, 2018).

Il ne s'agit pas en effet de « jeter le bébé avec l'eau du bain » : les données d'enquêtes par sondage demeurent toujours nécessaires, car elles

rendent compte d'autres dimensions de la réalité sociologique que les sciences des données ne peuvent prétendre décrire et expliquer. Elles permettent notamment de produire des enquêtes longitudinales auprès d'un même panel de répondants et répondantes ou d'approfondir des modèles explicatifs de causalité en mettant en avant des corrélations significatives entre les caractéristiques sociodémographiques classiques de la sociologie (âge, genre, niveau de scolarité, profession, lieu de résidence, appartenance de classe ou de groupe, position sociale, niveau de revenu) et les données tirées des enquêtes.

La multiplication exponentielle des données numériques : nouvelles perspectives pour les approches quantitatives

Même si l'expression *big data* a été largement utilisée depuis une décennie aussi bien dans les communautés scientifiques que dans la presse spécialisée, force est de constater que cette notion recouvre un grand nombre de pratiques diversifiées qui mobilisent des types hétéroclites de données quantitatives. Il y a un danger de confusion sémantique à vouloir amalgamer l'ensemble de ces pratiques hétérogènes sous un vocable unique. L'immense popularité de ce terme à la mode révèle pourtant une transformation sociétale et scientifique plus vaste, à savoir la multiplication exponentielle de données numériques, une explosion quantitative qui surgit sous l'effet combiné du développement d'Internet, de la multiplication des capteurs de traces numériques de toutes sortes et de l'informatisation généralisée des organisations (Ollion et Boelaert, 2015). Selon Ollion et Boelaert (*ibid.*, p. 302), ces données sont dites numériques d'abord en raison du fait qu'elles sont encodées sous forme de chiffres et sont interprétées par un processeur informatique, c'est-à-dire que cet encodage « permet à une machine de lire et d'exécuter des actions séquentielles sur ces informations ». D'où l'importance prise par les algorithmes dans la gestion contemporaine des organisations. Aussi, l'on ne peut nier qu'un tel saut quantitatif dans la grandeur des corpus de données numériques modifie significativement les pratiques de recherche, et en particulier, les approches quantitatives.

Les enjeux sociaux et scientifiques concernant la valorisation, la circulation et l'accès des données numériques se multiplient. De nouvelles compétences, de nouveaux domaines de formation universitaire, de

nouvelles spécialités scientifiques apparaissent – par exemple, les « sciences des données » (*data sciences*) se constituent, des revues scientifiques dédiées voient le jour (*Big Data & Society*) – et de nouveaux lieux industriels et scientifiques orientés vers l'analyse de grands corpus de données se créent (*datalabs, data centers*). De nombreux gouvernements nationaux, des centres de recherche interuniversitaires, de même que des institutions internationales ont réalisé que la production et la valorisation des données représentent une source importante de croissance et de développement économique. En même temps, cette explosion quantitative des données numériques fait surgir de multiples débats au regard des possibilités nouvelles s'offrant au domaine des méthodes en sciences humaines et sociales.

Ainsi en est-il des controverses et questionnements que suscitent les nouvelles possibilités méthodologiques associées à la multiplication des données numériques : l'hétérogénéité des informations peut-elle permettre l'universalité et la transposition immédiate d'une même méthode de traitement d'un corpus à un autre ? N'y a-t-il pas une pauvreté conceptuelle relative des données récoltées au moyen des techniques d'enregistrement automatique au regard de la pertinence scientifique de ces données quant aux questions précises de recherche qui sont en jeu ? Pouvons-nous surmonter les contraintes éthiques liées au traitement en masse de données personnelles, ces corpus pouvant par ailleurs être couplés avec des bases de données nominatives ? L'analyse de très grands corpus signifie-t-elle nécessairement l'addition d'un « supplément de connaissance » (Bastard *et al.*, 2017) aux problèmes sociologiques sous-jacents et qui étaient analysés jusqu'ici au moyen des approches quantitatives traditionnelles ? Il ne faut pas non plus perdre de vue le fait que de nombreux objets de recherche ne sont pas traitables par le recours aux approches quantitatives, ce qui laisse intacte la pertinence des approches qualitatives pour l'analyse de problèmes qui passent en dessous des radars de la quantification des traces, ce qui est certainement, comme nous le signalions précédemment, le fil conducteur du présent ouvrage.

Qu'en est-il finalement des innovations significatives en matière d'approches quantitatives ? Comme nous le disions, l'une des avancées incontestables consiste à pouvoir considérer les données de populations entières plutôt que l'analyse des performances de seuls échantillons, ce qui affecte le problème de la représentativité statistique. Dans ce contexte,

à côté des statistiques descriptives ou inférentielles classiques, un nouveau domaine méthodologique s'ajoute, celui qui s'appuie sur les méthodes d'apprentissage automatique (*machine learning*) issues de l'intelligence artificielle. Ces méthodes d'apprentissage statistique – d'abord inventées pendant la décennie 1960, mais ayant connu d'importants développements à partir des années 1990, avec un intérêt marqué pour les algorithmes d'apprentissage – permettent de traiter des échantillons de très grande taille, voire des populations entières. Elles se retrouvent appliquées dans plusieurs domaines : reconnaissance vocale, programmation de véhicules autonomes, génomique ou robotique (Ollion et Boelaert, 2015). Les méthodes d'apprentissage statistique sont désormais utilisées en sciences sociales, en particulier pour l'analyse textuelle. Dans un travail récent de recension critique des nouvelles techniques d'analyse textuelle, Cointet et Parasio (2018) s'interrogent pour savoir dans quelle mesure ces approches peuvent constituer une ressource pertinente pour l'enquête sociologique. Ils signalent notamment l'usage intéressant des analyses des réseaux sémantiques dans des corpus textuels. Toutefois, ce serait la modélisation thématique (*topic models*) qui retiendrait surtout l'attention des spécialistes de sciences sociales intéressés aujourd'hui par l'innovation méthodologique fondée sur l'analyse des données massives (Blei, 2012).

Or, ces nouvelles approches inductives propres aux « sciences sociales computationnelles » (Lazer *et al.*, 2009) ne feraient-elles pas l'impasse sur le contexte de construction sociale et scientifique spécifique de production des données qui constituent le matériau de ces grands corpus ? Il y a longtemps que l'on sait, en effet, que les données ne sont jamais neutres, qu'il est nécessaire de pratiquer une critique épistémologique des sources d'où elles proviennent. C'est-à-dire de procéder à un travail d'explication du contexte spécifique de production de ces données mettant en évidence les biais méthodologiques et idéologiques, voire de l'emprise invisible du sens commun et des prénotions sur les chercheurs et les chercheuses pouvant avoir influencé ce contexte de production cognitive (Bourdieu, Chamboredon et Passeron, 1973). D'où la nécessité de rester « épistémologiquement vigilants » à l'égard des catégories analytiques utilisées dans les observations et raisonnements en statistiques computationnelles.

Une autre piste d'innovation méthodologique dans le contexte des potentialités du numérique, suivie en particulier dans plusieurs travaux en Europe, est celle de l'exploration des méthodes numériques (*digital*

methods). L'une des figures clés de ce nouveau courant de recherche – nettement en développement dans le monde anglo-saxon à l'échelle internationale – est le professeur Richard Rogers (2013), de l'Université d'Amsterdam. Il est l'animateur scientifique de la Digital Methods Initiative, dont le site Web renvoie à une équipe de l'université néerlandaise, mais aussi à un réseau international de recherche. Cette approche implique qu'il y a un intérêt à travailler avec des outils et des données « nativement numériques », c'est-à-dire qui sont immédiatement proposés aux usagers et usagères, et produits directement par l'univers numérique, par exemple l'hyperlien, le fil de discussion d'un forum ou le mot-clic (*hashtag*) de la plateforme Twitter. Pensons aussi aux ressources offertes par les différents dispositifs numériques du Web, les moteurs de recherche ou les plateformes numériques. Considérons par exemple les données produites par les *analytics* (dispositifs d'analyse) des diverses plateformes (Google, Apple, Twitter, etc.). Le projet épistémologique des méthodes numériques (*digital methods*) consiste à « réorienter » l'usage de ces outils et de ces données nativement numériques de manière à construire une nouvelle vision en sciences sociales qui soit en résonance avec les transformations sociales et politiques contemporaines. Au-delà d'une finalité strictement descriptive des usages d'Internet, l'analyse de ces données nativement numériques permet ainsi de nourrir une façon nouvelle de faire de la sociologie en mettant par exemple en relief et à nouveaux frais le traitement des controverses sociales sur la scène publique.

Le constat que l'on peut formuler devant l'explosion quantitative contemporaine entraînant la constitution de très grands corpus de données – et par conséquent l'élargissement de la palette des approches méthodologiques – est de reconnaître un déplacement épistémique dans les intérêts de la recherche en sciences de l'information et de la communication. Jusqu'ici, la primauté méthodologique dominante était orientée vers le développement de modèles de causalité et d'explication de la réalité sociale. Aujourd'hui, par contraste, nous remarquons un renouveau pour les approches centrées soit sur la stricte description des situations, soit sur des démarches de classification des données en vue de la mise en relief par un processus inductif de nouvelles catégories pour penser la sphère sociale. On ne peut que reconnaître l'existence d'une « pluralité de modes de connaissance » des actions humaines (traduction libre de Savage et Burrows, 2009, p. 769). Nous sommes devant une panoplie de grilles de

lecture de la réalité humaine qui, d'ailleurs, peuvent interagir entre elles. Cette pluralité et cette diversité des méthodes, plutôt que de s'affronter ou de se concurrencer, devraient permettre l'invention de dispositifs méthodologiques hybrides et combinatoires. En même temps, on ne peut faire l'impasse sur le fait que la critique du choix des méthodes est nécessairement et intimement articulée à la nature des problèmes et des questions de recherche. La bonne méthodologie est celle qui permet de construire les données qui correspondent le mieux aux questions de recherche auxquelles nous voulons répondre. La méthodologie ne peut être pensée en vase clos ; elle est nécessairement liée à une problématique et à une question de recherche.

Tout l'enjeu méthodologique contemporain consiste à pouvoir se tenir suffisamment à distance des firmes propriétaires des outils, des infrastructures et des plateformes numériques de manière à maintenir l'indépendance nécessaire à une production scientifique entièrement libre vis-à-vis des contraintes qui pourraient être imposées par l'économie de marché et le capitalisme informationnel.

Au fil de notre argumentaire, nous avons vu de quelles manières le domaine des méthodes quantitatives était profondément transformé par ce que Bastin et Tubaro (2018), ont appelé « le moment *big data* des sciences sociales ». En parallèle, alors même que certains prophètes de malheur allaient jusqu'à prédire sa disparition, le domaine des méthodes qualitatives semble moins perturbé par ces mouvements de transformations épistémiques liées à la multiplication exponentielle des données transactionnelles. En ce sens, nous pourrions dire que les méthodes qualitatives sont un incontournable des pratiques des sciences sociales. Voilà un enseignement à retenir du présent ouvrage, à savoir que les méthodes qualitatives permettent la compréhension et l'analyse de phénomènes peu numérisés ou devenus relativement invisibles en raison des stratégies de privatisation des données quantitatives exercées par les firmes propriétaires des plateformes. Un univers de compréhension qui reste inaccessible aux méthodes quantitatives même transformées par les données massives. Ce développement asymétrique des méthodes laisse ainsi place à de pertinentes et astucieuses innovations en matière d'approches mixtes pour aborder le champ des pratiques numériques. Voilà peut-être la contribution la plus importante du présent ouvrage.

Références

- BASTARD, Irène *et al.*, « Facebook, pour quoi faire ? Configurations d'activités et structures relationnelles », *Sociologie*, vol. 8, n° 1, 2017, p. 57-82.
- BASTIN, Gilles et Paola TUBARO, « Le moment *big data* des sciences sociales », *Revue française de sociologie*, vol. 59, n° 3, 2018, p. 375-394.
- BLEI, David M., « Probabilistic topic models », *Communications of the ACM*, vol. 55, n° 4, 2012, p. 77-84.
- BOURDIEU, Pierre, Jean-Claude CHAMBOREDON et Jean-Claude PASSERON, *Le métier de sociologue*, Mouton, 1973.
- COINTET, Jean-Philippe et Sylvain PARASIE, « Ce que le *big data* fait à l'analyse sociologique des textes. Un panorama critique des recherches contemporaines », *Revue française de sociologie*, vol. 59, n° 3, 2018, p. 533-557.
- LATOUR, Bruno, *Changer de société, refaire de la sociologie*, La Découverte, 2007a.
- LATOUR, Bruno, « Beware, your imagination leaves digital traces », *Times Higher Literary Supplement*, 2007b.
- LAZER, David *et al.*, « Computational social sciences », *Science*, vol. 323, n° 5915, 2009, p. 721-723.
- OLLION, Étienne et Julien BOELAERT, « Au-delà des *big data*. Les sciences sociales et la multiplication des données numériques », *Sociologie*, vol. 6, n° 3, 2015, p. 295-310.
- PROULX, Serge, « La critique du capitalisme cognitif », dans France AUBIN et Julien RUEFF (dir.), *Perspectives critiques en communication*, Presses de l'Université du Québec, 2016, p. 191-212.
- PROULX, Serge et Julien RUEFF, *Actualité des méthodes de recherche en sciences sociales sur les pratiques informationnelles*, Centre d'études sur les médias, 2018.
- ROGERS, Richard, *Digital Methods*, The MIT Press, 2013.
- SAVAGE, Mike et Roger BURROWS, « Some further reflections on the coming crisis of empirical sociology », *Sociology*, vol. 43, n° 4, 2009, p. 762-772.
- SAVAGE, Mike et Roger BURROWS, « The coming crisis of empirical sociology », *Sociology*, vol. 41, n° 5, 2007, p. 885-899.

Notices biographiques

Claire Balleys est sociologue et professeure à la Haute École de travail social de Genève. Elle est l'auteurice de *Grandir entre adolescents. À l'école et sur Internet* (Presses polytechniques et universitaires romandes, 2015). Ses plus récents travaux ont été publiés dans les revues *Questions de communication*, *Genre, sexualité & société*, *Media, Culture & Society*, ainsi que *Social Media + Society*.

Claudine Bonneau est professeure agrégée au Département d'analytique, opérations et technologies de l'information de l'École des sciences de la gestion à l'Université du Québec à Montréal (UQAM). Elle est responsable de l'axe de recherche Mutations du travail et de l'activité contributive du Laboratoire sur la communication et le numérique (LabCMO), ainsi que membre du Centre interuniversitaire de recherche sur la science et la technologie (CIRST). Ses travaux ont notamment été publiés dans les revues *Educational Technology Research and Development*, *Communication, Research & Practice* et *International Journal of Project Management*.

Dominique Boullier est professeur des universités en sociologie à l'Institut d'études politiques de Paris (Sciences Po). De 2015 à 2019, il a été professeur à l'École polytechnique fédérale de Lausanne où il a enseigné la User Experience (UE) et dirigé le Social Media Lab. Il a dirigé ou créé plusieurs laboratoires de recherche dont Costech (UT Compiègne, où il a été professeur de 1996 à 2005) et le User lab Lutin à la Cité des sciences de Paris (2004 à 2008). Il a publié de nombreux ouvrages dont *Sociologie du numérique* (Armand Colin, 2019).

Nathalie Casemajor est professeure au Centre Urbanisation Culture Société de l'Institut national de la recherche scientifique (INRS) à Montréal. Son travail porte sur le développement culturel, la mobilisation citoyenne et la culture numérique. Elle est codirectrice de l'Observatoire des médiations culturelles

(OMEC) et a coédité l'ouvrage *Expériences critiques de la médiation culturelle* (Presses de l'Université Laval, 2017). Ses travaux ont été publiés dans *Journal of Peer Production, Intermediality* et *Communication & Organisation*.

Johann Chautet est sociologue et chercheur au Laboratoire interdisciplinaire Solidarités, Sociétés, Territoires (LISST) du Centre national de la recherche scientifique (CNRS). Il s'intéresse à l'étude des usages des technologies de l'information et de la communication, ses derniers travaux ont été publiés dans *RESET, Mobile Media & Communication* et *Genre, sexualité et société*. Il a codirigé l'ouvrage *Critiques du numérique* (L'Harmattan, 2018).

Alexandre Coutant est professeur au Département de communication sociale et publique de l'UQAM et directeur du Centre de recherche sur la communication et la santé (ComSanté). Il est également responsable de l'axe de recherche Mutations des sociabilités et de l'agir politique du LabCMO. Il a récemment codirigé le numéro *Trajectoires professionnelles en communication: atypies, hybridités et temporalités* dans la revue *Communication & professionnalisation*.

Jean-Claude Domenget est maître de conférences habilité à diriger des recherches en sciences de l'information et de la communication à l'Université de Bourgogne Franche-Comté. Membre du laboratoire Elliadd et du Réseau sur la professionnalisation des communicateurs (RESIPROC), il est également corédacteur en chef de la revue *Communication et professionnalisation* et vice-président de la Commission recherche de la Société française des sciences de l'information et de la communication (SFSIC). Il a lancé le Groupe de réflexion sur l'éthique et le numérique en information-communication (GENIC).

Gabrielle Silva Mota Drumond est doctorante en communication à l'UQAM, ainsi qu'édimestre-présentatrice (médias sociaux) à ICI Acadie. Maîtresse en communication, concentration en médias socionumériques (UQAM), elle est assistante de recherche, journaliste et gestionnaire de médias sociaux. Elle a aussi été coordonnatrice du pôle montréalais du LabCMO.

Nina Duque est doctorante en communication et chargée de cours à l'UQAM. Boursière du Conseil de recherche en sciences humaines (CRSH), elle est membre du LabCMO et de la Chaire de recherche sur les usages des technologies numériques et les mutations de la communication.

Julien Figeac est chargé de recherche au CNRS en sociologie et en sciences de la communication. Il est membre du Laboratoire Interdisciplinaire Solidarités, Sociétés, Territoires (LISST) de l'Université de Toulouse - Jean Jaurès. Il a publié divers articles traitant des usages d'Internet et des réseaux socio-numériques, où il étudie notamment les sociabilités en ligne et les formes connectées de participation politique.

Nicole Gallant est politologue et professeure au Centre Urbanisation Culture Société de l'INRS. Elle est également directrice de la *Revue Jeunes et Société (RJS)* et membre de la Chaire-réseau de recherche sur la jeunesse au Québec (CRJ). Parmi diverses publications, elle a codirigé l'ouvrage *Les jeunes et l'action politique: participation, contestation, résistance* (Presses de l'Université Laval, 2016).

Christine Hine est professeure de sociologie à l'Université du Surrey au Royaume-Uni. Ses plus récents livres portent sur l'ethnographie du Web (Bloomsbury Publishing) et sur la recherche qualitative en contexte numérique (Oxford University Press). Elle est également l'auteur de nombreux articles publiés dans *The Sage Handbook of Online Research Methods*, ainsi que dans les revues *Information, Communication & Society* et *Science, Technology & Human Values*.

Katherine Labrecque est doctorante en sociologie de la jeunesse et du numérique à l'INRS. Elle est titulaire d'un baccalauréat et d'une maîtrise en anthropologie de l'Université Laval ainsi que d'une maîtrise en pratiques de recherche et action publique de l'INRS. Boursière du Conseil de recherches en sciences humaines (CRSH), elle coordonne à titre d'auxiliaire de recherche les projets du chantier numérique de la Chaire-réseau de recherche sur la jeunesse du Québec.

Guillaume Latzko-Toth est professeur titulaire au Département d'information et de communication de l'Université Laval, codirecteur du LabCMO et membre du CIRST. Il a publié dans les revues *Communication, Canadian Journal of Communication, Recherches sociographiques* et *Bulletin of Science, Technology & Society*. Il a contribué à de nombreux ouvrages collectifs dont le *Handbook of Social Media Research Methods* (SAGE Publications, 2017) et *digitalSTS: A Field Guide for Science & Technology Studies* (Princeton University Press, 2019).

Florence Millerand est professeure titulaire au Département de communication sociale et publique à l'UQAM et codirectrice du LabCMO. Elle est titulaire de la Chaire de recherche sur les usages des technologies numériques et les mutations de la communication, ainsi que membre du CIRST. Elle a publié dans les revues *Social Media + Society, Information, Communication & Society, Science, Technology & Human Values* et *Revue d'anthropologie des connaissances*. Elle a codirigé l'ouvrage *La reconfiguration du travail scientifique en biodiversité: pratiques amateurs et technologies numériques* (Presses de l'Université de Montréal, 2018).

Mélanie Millette est professeure agrégée au Département de communication sociale et publique à l'UQAM. Elle est membre du LabCMO, où elle dirige l'axe Méthodes de recherche sur les usages. En plus de ses contributions à des ouvrages collectifs (dont le *Handbook of Social Media Research Methods*, SAGE

Publications), elle a publié dans les revues *Social Media+Society* et *Communication et Organisation*, et a codirigé l'ouvrage *Médias sociaux. Enjeux pour la communication* aux Presses de l'Université du Québec.

David Myles est professeur associé au Département de sexologie de l'UQAM et postdoctorant à l'Institut genre, sexualité et études féministes de l'Université McGill. Il est membre du LabCMO, du CIRST, de la Chaire de recherche sur l'homophobie de l'UQAM et du Digital Media Research Centre (Queensland University of Technology). Ses travaux ont été publiés dans les revues *New Media & Society*, *Information, Communication & Society*, *Social Media+Society* et *Réseaux*.

Dominique Pasquier est sociologue, directrice de recherche émérite au CNRS et membre du Centre de recherche sur les liens sociaux (CERLIS) à l'Université Paris Descartes. Elle a publié *L'Internet des familles modestes: enquête dans la France rurale* (Presses des Mines, 2018).

Madeleine Pastinelli est ethnologue de formation, professeure titulaire au Département de sociologie de l'Université Laval et membre régulier du Centre de recherche Cultures – Arts – Sociétés (CELAT) et du LabCMO. Elle a publié plusieurs articles sur les enjeux méthodologiques et éthiques de l'enquête en contexte numérique. Elle a également publié *Des souris, des hommes et des femmes au village global* (Presses de l'Université Laval, 2007) et codirigé l'ouvrage *S'engager en ligne* (UNESCO, 2019).

Daniel Pélissier est docteur en Sciences de l'Information et de la Communication et membre du laboratoire IDETCOM (Institut du Droit de l'Espace des Territoires, de la Culture et de la Communication) de l'Université Toulouse 1 Capitole. Ses travaux portent sur la communication numérique des organisations et sont publiés dans les revues *Communication & Organisation*, *Communication & Professionnalisation*, *Gérer & Comprendre* ainsi que *Communication & Management*.

Serge Proulx est sociologue et professeur émérite à l'École des médias de l'UQAM. Il est également professeur associé à Télécom ParisTech, conseiller scientifique et cofondateur du LabCMO. Il est l'auteur de nombreux ouvrages et rapports de recherche, dont *Actualité des méthodes de recherche en sciences sociales sur les pratiques informationnelles* (Centre d'études sur les médias, 2018) et *La participation numérique: une injonction paradoxale* (Presses des Mines, 2020).

Sandrine Roginsky est professeure à l'École de communication de l'UCLouvain. Elle est directrice du Laboratoire d'analyse des systèmes de communication des organisations (LASCO) et membre associée au LabCMO. Ses plus récents travaux

ont été publiés dans *Journal of Applied Journalism and Media Studies*, *Recherches en Communication* et *Communication & professionnalisation*.

Christine Thoër est professeure titulaire au Département de communication sociale et publique à l'UQAM et membre du LabCMO. Elle est également membre du Groupe d'études et de recherches axées sur la communication internationale et interculturelle (GERACII). Ses récents travaux ont été publiés dans *Communiquer*, *Recherches féministes*, *Social Media + Society*, *Santé mentale au Québec*, *Enfances*, *Familles*, *Génération*s ainsi que *Santé Publique*.

Table des matières

Préface	7
<i>Dominique Pasquier</i>	
Remerciements	13
Introduction	15
<i>Mélanie Millette, Florence Millerand, David Myles, Guillaume Latzko-Toth</i>	

PREMIÈRE PARTIE

Les méthodes en contexte numérique

CHAPITRE 1	
La redistribution des méthodes de recherche en contexte numérique : critique d'une cartographie	25
<i>Florence Millerand, David Myles, Serge Proulx</i>	
CHAPITRE 2	
Les traces numériques et le pouvoir d'agir des répliques	39
<i>Dominique Boullier</i>	
CHAPITRE 3	
L'analyse du numérique dans une perspective complexe avec la sociologie des usages	59
<i>Alexandre Coutant, Jean-Claude Domenget</i>	

DEUXIÈME PARTIE
Les méthodes ethnographiques

CHAPITRE 4

- L'ethnographie des communautés en ligne et des médias sociaux :
modalités, diversité, potentialités** 77
Christine Hine

CHAPITRE 5

- Les bénéfices mutuels de l'ethnographie et de l'analyse du discours
en contexte numérique** 103
David Myles

CHAPITRE 6

- Les terrains de recherche en ligne et hors-ligne : proposition
pour une triangulation des méthodes** 119
Sandrine Roginsky

CHAPITRE 7

- Les pratiques de visionnement connecté des jeunes :
méthodes traditionnelles, virtuelles et numériques** 137
Christine Thoër, Florence Millerand, Nina Duque

CHAPITRE 8

- Les usages numériques adolescents : vers une ethnographie multisite
de la « chambre numérique »** 153
Nina Duque

CHAPITRE 9

- L'observation des représentations adolescentes de l'intimité :
enquête sur YouTube** 165
Claire Balleys

TROISIÈME PARTIE
Les méthodes sur traces

CHAPITRE 10

- La densification des données : revaloriser la recherche qualitative
à l'ère des données massives** 181
*Guillaume Latzko-Toth, Claudine Bonneau,
Mélanie Millette*

CHAPITRE 11

**La visite commentée : documenter les pratiques numériques
par l'entretien sur traces** 195

*Nicole Gallant, Katherine Labrecque, Guillaume Latzko-Toth,
Madeleine Pastinelli*

CHAPITRE 12

**L'entretien d'autoconfrontation : analyse d'une activité de navigation
sur Internet** 211

Daniel Pélissier

CHAPITRE 13

**La collecte manuelle des traces d'usage par la découverte
progressive de mots-clics** 225

Claudine Bonneau

QUATRIÈME PARTIE

**Les méthodes d'analyse d'images
et les dispositifs vidéo**

CHAPITRE 14

**L'analyse visuelle du Web : une approche forensique
des métadonnées d'images** 243

Nathalie Casemajor

CHAPITRE 15

**Les contextes d'usage de la téléphonie mobile :
données fines et enregistrement vidéo** 259

Julien Figeac, Johann Chaulet

CHAPITRE 16

**La boîte noire de Netflix : la configuration des usages
par les interfaces et les algorithmes** 275

Gabrielle Silva Mota Drumond

POSTFACE

**Une nécessaire politique des méthodes pour la transition
numérique** 291

Serge Proulx

Notices biographiques 303



Ce livre fait le point sur les avancées et les enjeux de la méthodologie qualitative en contexte numérique. Il offre un panorama des méthodes de recherche les plus récentes pour l'étude des phénomènes en ligne par les sciences sociales – communication, sociologie, anthropologie. Cet ouvrage foisonnant présente des textes allant des fondements épistémologiques aux méthodes visuelles pour l'étude d'objets numériques, en passant par les stratégies ethnographiques, la collecte et l'analyse de traces d'activités en ligne. Au-delà du manuel, l'ouvrage se distingue par son actualité, sa cohérence entre la théorie et la pratique ainsi que par sa résonance interdisciplinaire.

Les chercheurs et les chercheuses en sciences humaines et sociales, les spécialistes d'Internet ainsi que les étudiants et les étudiantes intéressés à l'analyse d'un objet ou d'un fait social relié à l'univers numérique seront comblés notamment par la richesse des thèmes abordés, les types de corpus étudiés et les exemples d'enquêtes.

Mélanie Millette est professeure agrégée au Département de communication sociale et publique de l'UQAM.

Florence Millerand est professeure titulaire au Département de communication sociale et publique de l'UQAM.

David Myles est professeur associé au Département de sexologie de l'UQAM.

Guillaume Latzko-Toth est professeur titulaire au Département d'information et de communication de l'Université Laval.

Avec les contributions de Claire Balleys, Dominique Boullier, Claudine Bonneau, Johann Chaulet, Nathalie Casemajor, Alexandre Coutant, Jean-Claude Domenget, Nina Duque, Julien Figeac, Nicole Gallant, Christine Hine, Katherine Labrecque, Guillaume Latzko-Toth, Florence Millerand, Mélanie Millette, David Myles, Dominique Pasquier, Daniel Pélissier, Madeleine Pastinelli, Serge Proulx, Sandrine Roginsky, Gabrielle Silva Mota Drumond et Christine Thoër.

34,95 \$ • 31 €

Couverture : © Rawpixel.com

Disponible en version numérique
www.pum.umontreal.ca

ISBN 978-2-7606-4249-2



9 782760 642492