

Introduction

La face cachée des données

We get passports, birth certificates, parking tickets; we fill in forms to apply for jobs, for insurance, for dental benefits, we are given grades, diplomas, degrees; we pay bills and taxes; we read and answer advertisements; we order from menus in restaurants, take a doctor's prescription to the drugstore, write letters to newspapers, we watch television, go to the movies and so on and so on. Our lives, to a more extensive degree than we care to think of, are infused with a process of inscription, producing printed or written traces or working from them.

(Smith, 1984, p. 59)

Nombreux sont celles et ceux qui voient dans notre époque les signes d'une révolution. Un bouleversement technique, économique, politique au centre duquel se trouve une entité présente dans toutes les bouches et sous toutes les plumes : la donnée. Mais qu'est-ce exactement qu'une donnée ? Difficile à dire. Qu'elle soit assimilée à une ressource naturelle¹ ou présentée comme le résultat d'un phénomène climatique², cette entité de toutes les promesses (et de tous les dangers pour les plus pessimistes) n'est quasiment jamais décrite dans des termes précis. C'est ce qu'explique Borgman dans un livre récent qu'elle a consacré aux *data*, petites ou grandes : personne n'est véritablement capable de s'accorder autour d'une définition claire, exhaustive, de ce qu'est une donnée (Borgman, 2015, p. 28-29). À ce flou s'ajoute un doute. Telle qu'elle est présentée par ses promoteurs, comme par ses critiques, la « révolution des données » touche aussi bien les sciences que le marketing, la gestion, la logistique, ou encore la politique. Est-ce bien nouveau ? Qu'y a-t-il de révolutionnaire dans le rôle crucial que jouent aujourd'hui les données dans ces domaines ? Comme le laisse entendre la liste que fait Dorothy Smith en exergue de cette introduction, cela ne fait-il

1 La métaphore la plus courante est sans aucun doute celle du pétrole, que l'on pouvait notamment trouver dans le numéro de *Wired* de juillet 2014 : <http://www.wired.com/insights/2014/07/data-new-oil-digital-economy/>

2 Dans ce cas, c'est généralement l'image du déluge qui prévaut, comme le titre un célèbre article de *The Economist* : <http://www.economist.com/node/15579717>

pas déjà longtemps que nous vivons entourés de choses qui ressemblent à des données, y compris dans ce que nos vies ont de plus ordinaire ? N'est-ce pas une caractéristique centrale des sociétés modernes que de s'appuyer sur des masses toujours plus grandes données ?

Je ne veux pas dans ce livre répondre terme à terme à ces questions qui sont, on l'aura compris, artificiellement polémiques. Elles sont utiles pourtant, puisqu'elles soulignent l'intérêt de faire un pas de côté face aux apologies les plus enthousiastes et aux critiques les plus angoissées de cette prétendue révolution des données. Plutôt que de faire comme si tout allait de soi dans ce paysage, il me semble que les sciences sociales, en particulier la sociologie, peuvent participer à déplacer le regard sur les données, notamment en prenant au sérieux leurs ambiguïtés et la déjà longue histoire de leur omniprésence. C'est un premier mouvement dans ce sens que je propose de faire dans ce livre en dessinant une ligne de recherche qui soit capable d'interroger les données dans ce qu'elles ont de moins flagrant et paradoxalement de plus ordinaire. Ce pas suppose que l'on accepte de ne pas seulement faire de la sociologie avec des données, mais de la sociologie des données. Laissez-moi, pour illustrer à la fois l'incongruité et l'utilité ce geste, vous raconter une histoire personnelle.

Mon père est mort en mars 2013. Je pourrais bien sûr mobiliser des images très différentes pour rendre compte des semaines qui ont suivi la découverte de son corps. Parmi elles, il en est une qui donne à voir un aspect rarement évoqué pour décrire les temps du deuil, qui constitue pourtant une part importante de leur texture pratique. La mort plonge littéralement les vivants dans l'écriture et les transforme en producteurs compulsifs de données. Ainsi, il n'est pas exagéré d'affirmer que ma sœur, mes frères et moi avons traversé durant cette période un océan de mots, de chiffres, de formules. Un océan si vaste qu'il est impossible, trois ans plus tard, d'en recenser les détails. Même les archives que nous avons constituées en partie malgré nous n'y suffiraient pas, bien trop lacunaires. Pour être capables d'en rendre compte en détail, il aurait fallu que nous prenions des notes, que nous produisions nous-mêmes d'autres données, afin de garder en mémoire tout ce qui aujourd'hui est effacé, toutes les lettres, les documents, les formulaires que nous n'avons pas jugé utile de conserver. Toujours est-il que pendant plusieurs semaines, nous avons passé une partie de nos journées à lire, écrire, remplir des fiches, signer des documents, cocher des cases sur des sites Web, puis recommencer. Dans cette période, la mort de notre père, dont on pourrait imaginer qu'elle constituait un événement clairement identifié, immédiatement traduisible dans une information univoque se déplaçant sans heurts dans les canaux de communication de l'administration et des entreprises concernées – autrement dit une simple *donnée*, sa mort donc a dû sans cesse être reproduite sous des formes graphiques variées qui, bien que normalisées, n'étaient jamais complètement transparentes.

Pourquoi s'intéresser à ces données, aux activités d'écriture et de lecture qui accompagnent leur circulation, aux instruments qui permettent leur fabrication ? Si j'ai évidemment été personnellement affecté par tout cela, si ces opérations ont fortement marqué l'expérience très concrète de mon deuil, il n'y a peut-être pas de raison d'en faire un objet de recherche. En tant que sociologue, je devrais en fait pouvoir balayer d'un revers de la main ces artefacts informationnels et les basses besognes qui consistent à produire et à faire circuler des données administratives et commerciales. Après tout, ces documents ne sont-ils pas accessoires ? De simples moyens de communication qui n'ont pas grand-chose à voir avec le social dans sa complexité ni avec la réalité anthropologique de la mort. Les formulaires, les documents administratifs, les lettres d'attestation : autant d'ingrédients d'une paperasse qui constitue au mieux un « contexte » communicationnel encombrant, au pire les symptômes d'une bureaucratie qu'il faut critiquer en tant que système pris dans son ensemble. Certainement pas le cœur des véritables enjeux de nos sociétés contemporaines. Mais cette attitude de déni serait délicate à assumer jusqu'au bout tant l'occasion même d'un décès, précisément, rend marquante l'expérience des flux d'information qui irriguent notre vie en société. À moins d'adopter face à eux une position de surplomb qui donnerait au sociologue le rôle de trier à l'avance ce qui relève du social et ce qui n'en relève pas, les documents qui accompagnent la mort permettent de se faire une idée assez claire du caractère performatif de certaines informations. Ils font comprendre que la production des données, et les activités de lecture et d'écriture qu'elle suppose, participent pleinement à la fabrique de la réalité que nous partageons.

Difficile en effet de nier que ces données agissent. S'il est peut-être plus aisé dans le cas d'une naissance de réduire les formulaires que l'on remplit et les lettres que l'on envoie à des objets symboliques qui n'ont que peu de valeur face au bonheur d'accueillir un enfant, la force des données administratives se fait nettement sentir à l'occasion d'un décès. Dès les premières heures qui suivent le constat de la disparition d'un être proche, on est en effet confronté à l'écart qui existe entre cette information personnelle, située, partagée par quelques personnes dont le cercle s'élargit au fil des premiers appels téléphoniques, et la prise en compte effective de cette disparition au-delà de ce cercle. Il faut d'innombrables opérations informationnelles pour que la mort d'un proche ait lieu à l'échelle de la « société ». Il faut notamment que plusieurs institutions fassent concrètement disparaître de chacun de leurs fichiers, de chacune de leurs bases de données, de chacun de leurs dossiers clients les références à cet être. Dans d'autres cas, il faut au contraire qu'elles y fassent apparaître de nouvelles marques, qui indiquent son décès³. C'est seulement en passant par de denses réseaux informationnels que chaque mort peut ainsi être actée, devenir une « donnée » au sein d'instances pour lesquelles, tant que rien n'est fait, la personne demeure vivante sous des formes graphiques variées.

3 Sur le marché des activités liées à la mort et notamment sur les circuits administratifs des décès, voir le travail passionnant de Trompette (2008).

Voilà ce que je découvrais avec mes frères et ma sœur durant cette période de deuil : un à un, nous devions détricoter les fils qui composaient la trame des informations dont nous comprenions au fur et à mesure de nos demandes qu'elle maintenait notre père en vie à de nombreux points du monde. Pour sa banque, pour l'opérateur qui fournissait l'électricité de son appartement, pour celui qui le raccordait à Internet, sa mort n'avait pas l'évidence physique qu'elle avait pour nous.

Qu'on ne s'y trompe pas, il n'y a rien de métaphorique ou de symbolique dans cette affirmation. L'enjeu était on ne peut plus pratique : tant que nous n'agissions pas par écrit, les prélèvements automatiques continuaient de ponctionner son compte bancaire, le courrier ne cessait pas d'arriver dans sa boîte aux lettres (y compris ses factures de gaz, de téléphone, ses abonnements à la presse), et sa page Facebook affichait toujours les nouveaux messages de ses amis. En aboutissant au fur et à mesure de nos opérations dans des organisations très variées, nous comprenions que le réseau de données au sein duquel nous circulions faisait tenir l'identité de notre père dans ces institutions, ou plutôt la multitude de ses identités : locataire, contribuable, usager, client... Les documents que nous rédigeons, les formulaires que nous remplissons, les cases que nous cochons permettaient de tarir une à une les sources de ces identités.

En tant que sociologue, j'ai donc finalement quelques raisons de prendre au sérieux ces flux très concrets d'information, bien au-delà de la fascination pour une soi-disant « révolution des données », et de reconnaître que s'ils ne sont pas activés ou si certains des documents qui les traversent sont manquants, mal renseignés, quelque chose de l'ordre du « social » ne passe pas. Mais il se trouve qu'à l'époque où j'ai commencé à m'intéresser à cet aspect de la vie des données, j'étais en poste dans une école qui formait des ingénieurs de haut niveau dans le domaine des télécommunications. Et si j'avais évoqué mes questions auprès de mes collègues chercheurs ou de mes élèves, ceux-ci n'auraient sans doute pas bien compris la portée du problème. Dans les termes de l'ingénierie, l'enjeu est en effet assez simple. Sans cynisme aucun, on peut s'accorder pour voir dans la mort une information somme toute basique, sans ambiguïté : une donnée brute, qui n'a besoin d'aucun complément pour être enregistrée et diffusée. À l'ère de l'informatisation massive des administrations et des entreprises, et qui plus est dans le monde du *big data*, la transmission de cette toute petite donnée ne devrait poser aucun problème. Peut-être même que certains élèves fraîchement sortis de leurs classes préparatoires pourraient-ils imaginer qu'une donnée de ce genre peut aisément circuler par des mécanismes totalement automatisés, se diffusant à travers le monde à partir du premier signal que déclencherait le constat du décès. Défini abstraitement, le problème de l'extinction des identités n'a en effet rien de compliqué et, *sur le papier*, la circulation des données administratives ne présente aucun intérêt en soi ; il n'y a vraiment pas de quoi en faire un objet de recherche.

Une fois encore, il faut faire un pas de côté pour répondre à ces réticences et prendre la mesure du phénomène par-delà – ou en deçà – de sa version désincarnée. Le mépris de l'ingénieur pour la donnée administrative en tant que telle est semblable à celui du sociologue qui ne voit dans les activités qui la font exister et circuler que de la paperasse sans réelle importance : tous deux reposent sur le refus d'appréhender ce monde des données de plus près, d'en explorer l'épaisseur, et de prendre en considération l'expérience de celles et ceux qui y prennent part. Or, si je reviens à ce que j'ai vécu durant les semaines qui ont suivi la disparition de mon père, de même que je peux difficilement nier la force performative des données administratives que j'ai participé à fabriquer et à faire circuler, je ne peux pas non plus prétendre que l'ensemble de cette production informationnelle s'est réalisé sans problème. Au modèle abstrait d'une donnée qui se diffuserait automatiquement, je dois opposer les erreurs, les redondances, les incompréhensions et les ratés qui ont ponctué notre périple commercial et administratif. Même lorsque nos demandes « fonctionnaient » et que nos démarches aboutissaient, nous étions bien loin de la transparence et de la fluidité informationnelle auxquelles renvoie le vocabulaire de la transmission, si cher aux ingénieurs⁴.

Quelques exemples suffiront à donner une idée de l'épaisseur pratique de cette difficile production informationnelle. Commençons par la banque. C'est la première institution à laquelle mes frères, ma sœur et moi avons eu affaire. Grâce à des personnes compréhensives qui ont accepté d'interroger l'état des comptes d'un de leur client, l'absence totale de mouvement financier depuis plusieurs jours (hormis les prélèvements automatiques) a pu être avérée. Ce premier contact avec les réseaux des données de l'identité paternelle a déclenché les opérations qui ont abouti au constat officiel de son décès. Loin d'être le fruit d'un échange transparent d'informations factuelles, il n'a été possible qu'au prix de contournements de règlements et de procédures, d'ajustements dans les circuits standard des données bancaires. Mes frères, qui se sont présentés les premiers à l'agence locale, n'avaient rien d'autre que leur pièce d'identité à fournir en témoignage de leur bonne foi ; pas même un numéro de compte, rien qui les reliait immédiatement à notre père si ce n'est leur nom propre, fort répandu par ailleurs. Une fois confirmé que notre père était client de l'établissement, ils ont bénéficié de la bienveillance de l'agent au guichet et de sa supérieure hiérarchique, qui ont accepté de consulter les comptes en question, sans leur montrer directement, et de partager leur constat quant à l'inactivité prolongée de la carte bancaire.

4 Sur le frontispice de Telecom ParisTech, l'école historique des Postes et Télécommunications au sein de laquelle j'ai eu le plaisir d'enseigner et de développer mes travaux de recherche pendant plus de dix ans, on trouve une sculpture, accompagnée d'une légende qui résume parfaitement cette idée : « L'homme au cours des âges utilise les forces élémentaires pour les transmissions ». Et son laboratoire de recherche, auquel j'étais rattaché jusqu'en 2014, avait pour nom « Laboratoire de Transmission et de Communication de l'Information. »

Plus tard, nos relations avec cette banque changeront du tout au tout. Une fois le décès officialisé, les mêmes personnes avec lesquelles mes frères avaient échangé au guichet ont refusé catégoriquement de nous recevoir. Le dossier client dépendait maintenant d'un autre service, avec lequel nos interactions, nettement plus houleuses, ont pris une tournure beaucoup plus normalisée. Pour accéder aux comptes de notre père, y compris pour demander leur clôture, il nous fallait désormais produire plusieurs pièces officielles, et remplir de nombreux formulaires. Il fallait aussi payer certaines sommes forfaitaires, prix de l'impression des archives, entre autres services bancaires. Bon an mal an, après une semaine d'envois de messages électroniques et de courriers postaux, nous avons pu faire bloquer les prélèvements sur ces comptes, transférer certains droits à notre notaire et confirmer qu'aucun jeu d'écriture problématique n'avait eu lieu dans les derniers mois.

Bien plus compliquée a été la gestion des crédits qu'avait contractés notre père, et surtout celle de leurs assurances. Le décès du point de vue des organisations mandatées pour ce type d'assurance n'est pas en effet une affaire si univoque que j'ai bien voulu le faire croire plus haut. L'une des questions à laquelle celles-ci cherchent à répondre face aux éventuels héritiers (qui peuvent refuser la totalité du leg si des crédits trop importants restent à leur charge) n'est pas «Notre client est-il mort ou vivant?», mais «De quoi est-il mort?» Or, cette question n'est pas si simple qu'il n'y paraît. Elle s'appuie sur un document qui a été central tout au long du processus : le certificat de décès, pièce que nous leur avons fait parvenir aux premiers jours de nos démarches.

Une parenthèse s'impose ici, pour balayer une fois pour toutes les certitudes de l'ingénieur qui sommeille en chacun. Ma famille et moi savions que ce document comportait des informations que nous pourrions qualifier d'erronées. La date du décès notamment était inexacte, le médecin qui était de service ce jour-là ayant simplement reporté sur le formulaire la date de la découverte du corps. Nous avons compris à cette occasion que cette date était systématiquement «fausse» dans les situations comme la nôtre. À moins qu'il y ait des témoins, ou que les conditions de la mort nécessitent une enquête judiciaire qui exigerait que le jour et l'heure précise du décès soient connus, une mort est généralement antidatée sur le document on ne peut plus officiel qui l'atteste, pour la simple raison qu'il est impossible à partir de la seule situation du constat, sans examen approfondi, d'établir la date d'un décès, une donnée qui n'est pas si «donnée» que ça. Fin de la parenthèse.

S'il est légalement crucial, ce certificat ne représentait en réalité qu'une pièce parmi beaucoup d'autres aux yeux des personnes embauchées par les compagnies d'assurance pour instruire le cas de notre père. Entre autres informations, il a fallu notamment que nous reconstruisions ses antécédents médicaux, afin de leur faire parvenir dans une forme recevable. Cette étape était particulièrement complexe.

Outre ses propres archives dans lesquelles nous avons cherché des documents pouvant témoigner de sa santé, nous avons dû entrer en contact avec plusieurs médecins dont certains nous opposaient le secret professionnel tandis que d'autres nous signifiaient simplement qu'ils n'avaient pas le temps de s'occuper de notre cas. Nous échangeons par ailleurs régulièrement avec différents représentants des assurances au téléphone, qui mettaient à jour le dossier informatique au fil de nos interactions, ou qui nous annonçaient qu'il manquait encore une pièce « dans le système » et qu'il fallait tout faire pour leur envoyer.

Peu à peu, nous avons compris que ça n'était pas seulement à la question « De quoi est mort notre client ? » qu'il nous fallait répondre. Il y avait autre chose en jeu. Nous réalisons qu'en montant ce dossier, nous nous attelions en fait à la déconstruction d'une autre opération scripturale, effectuée par notre père quelques années auparavant. En souscrivant aux crédits concernés, celui-ci avait en effet lui-même coché des cases et signé des formulaires par lesquels il attestait qu'il ne souffrait pas de maladies graves qui auraient empêché l'obtention d'une assurance. En rassemblant les éléments que l'on nous demandait, nous procédions de fait à la vérification de ces données, au dépliage de ces signatures et de ces cases cochées quelques années plus tôt. Cette opération permettait finalement de répondre à une autre question, plus délicate encore : « A-t-il menti ? » Maintenant que les compagnies d'assurance étaient susceptibles de couvrir ces crédits, elles mettaient à l'épreuve ce maillon dans le réseau de données qui avait permis initialement de les accorder. Dans cette épreuve, nous étions chargés d'en démontrer la solidité en le complétant d'une autre branche. Or, cette consolidation n'avait rien d'évident ni d'automatique. Jusqu'à ce que l'annonce officielle de l'acceptation du dossier nous parvienne par le biais d'un appel téléphonique qui fut suivi d'une lettre quelques jours plus tard, nous avons navigué à vue, à la fois pour identifier les types de documents que nous devons fournir, mais aussi pour comprendre précisément ce qu'attendait chaque compagnie d'assurance dont les employés que nous contactions changeaient d'un appel à l'autre et paraissaient parfois eux-mêmes douter de l'état d'avancement de notre dossier, voire des éléments que nous devons procurer pour le faire progresser. Cette annonce finale, heureuse, nous a permis de confirmer notre décision d'accepter l'héritage et a ouvert la voie à d'autres activités scripturales à propos desquelles je ne m'étendrai pas ici, qui, elles-mêmes, ont abouti à l'élaboration d'un document à nouveau crucial : le certificat d'hérité.

Le certificat de décès, quant à lui, a continué d'être répliqué durant cette période. Il était le point de départ de chaque opération par laquelle nous tâchions de désactiver les services les uns après les autres. Dans la plupart des cas, moyennant quelques ajustements, ces opérations demeuraient relativement simples. Les abonnements à l'électricité, au gaz, à Internet n'ont pas posé trop de problèmes et nous comprenions que nous avions affaire à des institutions au sein desquelles notre situation avait toute sa place. C'est tout juste s'il n'y avait pas dans certains cas un numéro de touche à taper pour basculer, dès l'étape du tri mécanique de nos appels,

vers le service compétent. Elles n'étaient toutefois jamais totalement transparentes ni automatisées. Chaque dossier était accompagné d'échanges téléphoniques, et d'une enquête, même minime, à l'issue de laquelle nous finissions par remplir correctement les formulaires et comprendre les attentes précises de chaque institution. Jamais les données qu'il nous fallait faire circuler n'apparaissaient sous la forme d'un flux qu'il aurait suffi de diriger dans la bonne direction.

Notre dernière inquiétude concernait le compte Facebook de notre père. Comment parvenir à le fermer ? Il nous a fallu quelques recherches pour essayer d'y voir clair dans les procédures qui avaient été inventées par l'entreprise pour traiter ce genre de cas. Il était évident que celle-ci ne considérait pas la mort de ses clients tout à fait de la même manière que celles, bien plus anciennes, auxquelles nous avons eu affaire jusque-là. Sans doute d'ailleurs que les règles restaient floues, changeantes, d'autant plus qu'il s'agissait d'une multinationale dont l'ancrage juridique ne nous apparaissait pas clair. Nous avons finalement pu éviter la confrontation avec les employés de Facebook en découvrant par déduction le mot de passe du compte de notre père et en procédant nous-mêmes à sa fermeture.

Voici, rapidement esquissées, quelques lignes qu'il faudrait suivre à la loupe pour véritablement donner à voir la complexité des opérations informationnelles grâce auxquelles ma sœur, mes frères et moi-même avons progressivement « actualisé » la mort de notre père en produisant et en faisant circuler, parfois difficilement, un nombre considérable de données. À partir d'un exemple simple, elles donnent une première idée de l'importance de ces données et de l'intérêt de se pencher sur leur vie ordinaire, bien au-delà des seules obsessions actuelles pour le « numérique » et le tsunami informationnel qui caractériserait notre époque.

Mais comment parler de ces données ordinaires, sous quel angle les interroger ? Les choix narratifs que j'ai faits jusqu'ici pointent de manière évidente vers deux directions. La première mène aux formulaires et plus généralement aux instruments d'écriture qui cadrent la circulation des informations. Pour décrire le rôle des données dans notre monde, mettre en lumière leur force et le type d'actions qu'elles accomplissent, on pourrait en effet se préoccuper de la rigidité graphique et matérielle des artefacts qui les font exister, en s'inspirant par exemple de Michel Foucault et en étudiant chaque formulaire administratif comme un microdispositif de pouvoir, un outil essentiel des formes de gouvernementalité contemporaines, dont on sait qu'elles s'appuient sur un nombre toujours plus croissant de données. On tenterait alors de déceler à même les instruments d'écriture des traces de cette gouvernementalité, et de comprendre comment ils participent à produire les objets et les sujets qui sont ainsi gouvernés. En parallèle, cette perspective permettrait de questionner sous un angle original le matériau même sur lequel reposent les vastes dispositifs de calcul mis en œuvre par l'État et les entreprises commerciales qui manient ces instruments. Elle rendrait aux données leur forme initiale.

La seconde piste renvoie à l'expérience personnelle que l'on peut faire de ces données. En partant de l'exemple du décès à travers lequel j'ai dépeint des opérations parfois déplaisantes, voire douloureuses, il pourrait être intéressant d'élargir le champ de la description et de montrer que la production des données, le renseignement des cases et des « blancs » à remplir n'est pas toujours une affaire aisée, qui se réglerait sans heurts. En soulignant le caractère pesant de ces formalités, on obtiendrait ainsi une vue complémentaire sur les enjeux de pouvoir que les données cristallisent. On pourrait notamment prendre la mesure du coût que représente la rigidité des instruments d'écriture du point de vue de celles et ceux qui doivent s'y plier pour s'assurer que leur demande soit traitée ou que leur cas soit pris en compte. C'est bien ce sur quoi j'ai insisté en racontant cette histoire. Il nous a fallu nous ajuster et travailler à la fabrique même des données qui devaient inscrire le décès de notre père dans les différentes institutions auxquelles nous avons affaire. Le prix que l'utilisateur ou le consommateur doit payer pour produire des données, voilà une question de plus en plus importante à l'heure où de nombreuses entreprises semblent tirer une valeur considérable de tâches accomplies par leurs clients.

Les instruments informationnels eux-mêmes et les opérations que leurs principaux utilisateurs sont appelés à réaliser sont donc deux horizons vers lesquels pointe l'exemple avec lequel j'ai décidé d'ouvrir ce livre. Ces directions sont tout à fait pertinentes et seraient toutes les deux passionnantes à suivre. Plusieurs auteurs les ont d'ailleurs déjà empruntées, avec succès. Elles laissent toutefois dans l'ombre une autre histoire possible, une face cachée des données : celle du travail dont elles font l'objet dans les coulisses de leurs usages. Dans le récit que j'ai esquissé, ce travail, tout comme celles et ceux qui l'accomplissent, est quasiment invisible. Ça n'est en tout cas pas de leur point de vue que j'ai pu rendre compte de cette histoire. Pourtant, au-delà même des personnes que nous avons directement côtoyées par téléphone ou en face à face, il ne fait aucun doute que de nombreux employés ont examiné, agencé, et vraisemblablement ajusté, corrigé, les informations qui ont circulé à l'occasion du décès de notre père. Leurs activités, qui se sont réalisées en partie avec nous, parfois malgré nous voire contre nous, sont évidemment vitales aux flux d'informations que j'ai évoqués. Ces travailleuses et ces travailleurs en *back office*, que nous avons appelés il y a quelques années avec David Pontille les « petites mains de l'information » (Denis et Pontille, 2012c), traitent, vérifient, transforment, organisent les données. Autant d'opérations qui sont essentielles au fonctionnement de la mécanique informationnelle et de l'infrastructure qui, dans ce cas précis, assurait à mon père une part non négligeable de son existence administrative et commerciale, et qu'il a fallu « couper » brin après brin, jusqu'à la réduire aux seules données qui instauraient son décès, à celles qui attestaient de sa présence passée, et à celles qui désignaient les liens auxquels nous, ses enfants, pouvions prétendre.

Peu de chercheurs se sont penchés sur ces personnes et sur les conditions de leur travail⁵. Je voudrais inviter tout au long de ce livre à les étudier au plus près de leurs activités. L'enjeu principal de ce choix tient à l'invisibilité de ces opérations, qui semble grandissante, et qui doit être questionnée de front. L'augmentation considérable des services en ligne, les nombreuses possibilités qu'ils offrent quant à l'autonomie de leurs usagers qui, dans le même temps qu'ils bénéficient de fonctionnalités inédites, sont amenés à produire une part toujours plus importante de la relation de service, instaure par exemple de nouvelles formes d'invisibilité. Et le fait que cette tendance a principalement été étudiée comme une « mise au travail des clients » renforce encore ce qui peut apparaître comme une *back-officisation* d'une partie des activités essentielles au fonctionnement d'une société dite de l'information. De même, tandis que les données sont présentées comme une nouvelle ressource inestimable pour les entreprises et les administrations, avec l'avènement ces dernières années du domaine du *big data*, on ne peut qu'être frappé par le peu de questionnements, et de curiosité même, que suscitent les modalités de production de ces données désormais massives. Leur fabrication et les conditions de leur circulation restent dans l'angle mort de la plupart des réflexions, qui ne se focalisent que sur les formidables promesses des traitements dont peuvent faire l'objet ces données *a posteriori*, sur les connaissances qu'elles ne manqueront pas de faire émerger, ou à l'inverse sur les dangers que leurs usages représentent, notamment en termes de libertés individuelles.

Finalement, les idées mêmes de société de l'information ou de révolution des données, au sein desquelles l'immatérialité et la transparence sont revendiquées comme moteurs de progrès, participent de cette mise en invisibilité. Le vocabulaire qui leur est associé opère une véritable négation des basses œuvres que persistent à accomplir celles et ceux qui, même s'ils travaillent en grande partie sur écran, demeurent des « paperassiers » tenus dans l'ombre des données.

Voici donc le pas de côté que je souhaite effectuer dans ce livre : dessiner les grandes lignes d'une sociologie des coulisses de la donnée. En faisant cela, je ne fais pas grand-chose de plus que suivre les conseils de Leigh Star (1999), qui n'a eu de cesse de montrer l'utilité de « faire remonter à la surface le travail invisible » qui se niche dans l'ombre des situations les plus anodines. On l'aura compris, une telle posture permet notamment de contrebalancer les discours, des plus enthousiastes aux plus critiques, qui s'évertuent à présenter les données comme le moteur purement technique d'une révolution en marche. En explorant des activités de travail qui caractérisent une grande part de l'emploi contemporain, bien au-delà des pays industrialisés occidentaux, nous pouvons interroger sous un angle original les forces dont sont aujourd'hui dotées les données. Pour le dire

5 Je reviendrai en détails dans les prochaines pages sur les précieux travaux qui ont exploré cette part la plus ordinaire du travail administratif, en particulier ceux de JoAnne Yates, Delphine Gardey, ou encore Jean-Marc Weller.

en des termes théoriques, la plongée dans le travail invisible des petites mains de l'information vise à décrire un aspect mal connu des « conditions de félicité » de la performativité des données. Ce qui, dans leur existence la plus ordinaire, assure leur capacité à agir. Plus généralement, comme l'écrivait Delphine Gardey en ouverture de son histoire des employés de bureau, l'enjeu est de :

« rendre visible l'invisibilité du travail administratif, pour mieux comprendre de quelles activités sont faites nos économies contemporaines ou sur quelles infrastructures techniques et cognitives reposent les sociétés dites de l'information » (Gardey, 2001, p. 13).

Évidemment, je ne prétends pas offrir avec ce livre un panorama complet, qui traiterait de l'ensemble des aspects du travail invisible des données. On voit mal d'ailleurs à quoi pourrait bien ressembler un ouvrage répondant à une telle ambition, tant, précisément, ce qui relève de ce travail informationnel s'étend par delà les frontières géographiques et administratives. Bien plus modestement, je souhaite faire dans les pages qui suivent deux choses. Rappeler d'abord que les données ne sont pas des objets radicalement nouveaux et qu'elles ont donné lieu à des recherches passionnantes dans des domaines variés. Mon premier geste consistera donc à identifier ce que l'on sait déjà des données, de leur place dans la plupart des sociétés modernes, et bien sûr de leur travail au sens général du terme. Je montrerai à cette occasion qu'une sociologie du travail des données a tout à gagner à se construire dans la lignée de l'anthropologie de l'écriture, de la sociologie et de l'histoire des sciences et des techniques, et des études critiques de la comptabilité. Les cinq premiers chapitres de ce livre dessineront ainsi un double parcours. Un parcours historique, d'une part, qui permettra de mettre en perspective ce qui est généralement présenté comme une révolution contemporaine, et d'esquisser une première image des conditions concrètes du travail des données, telles qu'elles se sont progressivement stabilisées dans des configurations très variées. Un parcours analytique, d'autre part, au fil duquel seront articulées des traditions de recherche qui, si elles ne sont évidemment pas complètement étrangères, ne sont que rarement appréhendées dans un même mouvement, et dont les apports conceptuels sont précieux pour questionner les données et leur travail sans tomber dans les ornières des notions fourre-tout que semblent attirer ces objets depuis quelques années.

Dans un second temps, j'entrerai dans les détails de deux de mes propres enquêtes, qui me permettront de mettre en lumière deux aspects fondamentaux du travail des données contemporain : l'horizon toujours présent de la mécanisation, et donc de la dévalorisation, d'un grand nombre d'opérations qui se jouent dans les coulisses des bases de données commerciales et administratives ; le postulat très largement partagé d'une disponibilité universelle des données (qu'il faudrait simplement « libérer »), dont l'existence même n'a pourtant rien d'évident.

On l'aura compris, je ne prétends pas dans ce livre fournir une description exhaustive d'un phénomène. C'est, bien plus modestement, une série de questions que je chercherai à faire émerger et à affiner au fil des pages, des manières d'interroger les infrastructures informationnelles qui forment le socle de nos vies sociales, leur caractère à la fois ordonné et désordonné, de comprendre ce que leur fonctionnement quotidien doit à l'engagement d'innombrables travailleuses et travailleurs, d'explorer les processus qui mènent à l'invisibilité de ce travail, la variété de ces dynamiques d'invisibilisation, toujours relationnelles. Ces questions sont autant de détours nécessaires pour poser à nouveau, en des termes plus justes et plus concrets, celle qui a ouvert cette introduction : qu'est-ce qu'une donnée ?

Table des matières

PRÉFACE - PHYSICALITÉS DE L'ÈRE LIQUIDE : UNE SOCIOLOGIE	9
REMERCIEMENTS.....	13
INTRODUCTION - LA FACE CACHÉE DES DONNÉES.....	17
CHAPITRE 1 - DONNÉES EN LABORATOIRE : SUR LA PISTE DES INSCRIPTIONS.....	29
La production scripturale des faits scientifiques.....	29
La fabrique de textes transparents.....	35
Le laboratoire et ses inscriptions.....	37
Conclusions.....	44
CHAPITRE 2 - RAISON GRAPHIQUE ET MOBILES IMMUABLES.....	47
Écriture et posture critique.....	50
L'organisation spatiale de la cognition	53
De l'immuabilité à la mobilité	57
Conclusions.....	61
CHAPITRE 3 - ÉCRIRE, COMPTER, GOUVERNER, ORGANISER.....	63
Raison graphique, raison administrative.....	64
Formaliser, contrôler.....	69
Tenir à distance	76
Automatismes	78
Conclusions.....	80
CHAPITRE 4 - LA CONSTRUCTION SCRIPTURALE DE LA RÉALITÉ.....	83
Réalités documentaires	85
La politique des accounts.....	89
Conclusions.....	94
CHAPITRE 5 - INVERSIONS INFRASTRUCTURELLES	99
Une infrastructure scripturale	100
L'invisibilité relative des infrastructures.....	105
Ethnographie de l'activité, travail visible, travail invisible	107
Conclusions.....	111

CHAPITRE 6 - DANS LES COULISSES DU DOSSIER CLIENT	115
Rationalisation et investissement dans le «tout écrit»	117
Manipulations, transformations	122
Réduire les gouffres du dossier.....	131
Les données, les métadonnées et leur travail.....	141
Conclusions.....	146
CHAPITRE 7 - L'INSTAURATION DES DONNÉES	151
À la recherche des données de la cyclabilité	155
Hétérogénéité informationnelle	160
Des données, oui mais.....	167
Conclusions.....	177
CONCLUSIONS - CE QUE LE TRAVAIL DIT DES DONNÉES	181
RÉFÉRENCES BIBLIOGRAPHIQUES	185