N° 2 | 2021
Traces, données et preuves en contexte numérique : quelles acceptions interdisciplinaires ?

« Suivre à la trace », du numérique à l'éthique.

Quelle responsabilité pour un éditeur de logiciels ?

Lénaïk LEYOUDEC *Docteur*
TSH
COSTECH (EA 2223)
University of Technology of Compiègne
Marion GENAIVRE *Professionnel*
Agence de philosophie Thaé, Paris

Édition électronique :
URL :
https://intelligibilite-numerique.numerev.com/index.php/articles/revue-2/2151-suivre-a-la-trace-du-numerique-a-lethique
DOI : numerev_1706
Date de publication : 21/12/2021

CertiScience© *Certifié évalué par les pairs*

Cette publication est sous licence **CC BY-NC-ND** (Attribution - No commercial - No derivatives).

Pour **citer cette publication** : LEYOUDEC, L., GENAIVRE, M. (2021) « Suivre à la trace », du numérique à l'éthique.. *Intelligibilité du numérique*, (2). https://doi.org/10.34745/numerev_1706

Chaque organisation produit quotidiennement un nombre toujours croissant de données, formant les « *big data* ». Leur gestion soulève aussi bien des enjeux techniques, économiques, politiques qu'éthiques. Dans notre recherche, nous avons considéré le cas où un prestataire en édition logicielle est mandaté par une organisation pour y apporter une réponse technologique. Nous avons suivi les étapes de l'implémentation de la solution – une plateforme de réconciliation des données – et vérifier la promesse du prestataire. Depuis ce point d'observation, une question est apparue : une technologie capable de réconcilier des données doit-elle rendre intelligible toutes les données qu'elle peut légalement traiter et produire ? Le dilemme éthique soulevé a été appréhendé par le biais d'un référentiel éthique, dont les cinq principes ont servi de cadre à l'analyse d'un corpus documentaire construit autour de la plateforme et de son paratexte. Si l'éditeur logiciel semble en conformité avec le référentiel, un malaise éthique demeure et inspire deux recommandations – transparence et responsabilité – à l'intention des éditeurs logiciels « big data ».

Every organisation produces an ever-increasing amount of data on a daily basis, known as "big data". Their management raises technical, economic, political and ethical issues. In our research, we considered the case where a software publisher is mandated by an organisation to provide a technological response. We followed the implementation of the solution - a data reconciliation platform - and verified the provider's promise. From this vantage point, a question emerged: does a technology capable of reconciling data have to make intelligible all the data it can legally process and produce? The ethical dilemma raised was addressed by means of an ethical reference framework, whose five principles served as a frame for the analysis of a body of documentation built around the platform and its paratext. Although the software publisher seems to be in compliance with the reference framework, an ethical unease remains and inspires two recommendations - transparency and accountability - for "big data" software publishers.

Mots-clés :
Big data, Éthique, Responsabilité, Gouvernance des données, Plateforme, Big data, Ethics, Platform, Data governance, Accountability

---

1. Introduction

Dans l'« âge du pétabyte » (Anderson, 2008) dans lequel nous vivons, chaque organisation produit quotidiennement un nombre toujours croissant de données, faisant ainsi face à l'enjeu connu sous le nom de « *big data* », soit « la capacité à traiter en temps quasi réel d'immenses gisements d'informations, de grandes masses de données structurées, semi structurées, voire non structurées et disparates pour extraire de leur recoupement de nouvelles valeurs qui demeuraient jusqu'alors inconnues » (Zolynski, 2015, p. 25).

Mais qu'est-ce qu'une donnée au juste ? Plus qu'une trace et moins qu'une preuve ? Quel continuum y a-t-il de l'une à l'autre ? Et si trace, donnée et preuve n'ont pas le même statut, qu'est-ce qui les distingue ? Afin d'éclairer notre terrain de manière pertinente, un point conceptuel s'impose. Pour ce faire, nous avons choisi d'adopter une posture épistémique interdisciplinaire, qui croise l'apport des sciences de l'information et de la communication et celui de la philosophie[1].

Si la notion de donnée s'entend aujourd'hui principalement en son sens informatique - la « représentation d'une information sous une forme conventionnelle adaptée à son exploitation » (Académie française, 2005) –, nous devons essayer de garder à l'esprit ce que le terme nous suggère de manière plus littérale : une donnée est un don, ce qui est *là*, au sens de ce qui n'est pas cherché et donc en quelque sorte s'impose. Et c'est précisément ce qui lui permet de servir de socle à une réflexion partagée. La donnée désigne du connu sur lequel on peut s'appuyer pour aller vers l'inconnu. Mais le connu suppose néanmoins un accord, une reconnaissance, un régime d'évidence ou de convention qui fait que l'ensemble des participants à une réflexion s'entendent sur « ce qui est donné ».

Est-ce à dire que donnée vaut information ? Pas selon le philosophe Sven Ove Hansson, pour qui « il est regrettable qu'en informatique « savoir » et « information » soient souvent traités comme des synonymes. » (Hansson, 2002, p. 6) L'auteur précise :

> « *Les données diffèrent de l'information en ce qu'elles n'ont pas à se présenter sous une forme qui se prête à l'assimilation. Si au lieu de l'ouvrage en question, j'avais sur mon bureau les dix mille questionnaires sur lesquels il repose, j'aurais des données au lieu d'information. En résumé, il faut que des données soient assimilables pour pouvoir constituer de l'information et qu'elles soient assimilées pour pouvoir constituer du savoir.* »

L'information consiste donc en des données organisées qui, placées dans un contexte, sont dotées de sens. Si la donnée est une sorte de traduction la plus immédiate, la plus brute, d'un fait (l'unité minimale d'observation qui permet de le caractériser), on ne peut prétendre à sa totale objectivité. Il y a toujours de l'intention et du projet dans une donnée, ne serait-ce que parce que sa recherche procède d'une discrimination : vous

choisissez de mesurer telle variable plutôt que telle autre, et par ce filtre vous définissez une réalité que vous avez besoin de connaître.

Qu'une donnée soit le fruit d'une intentionnalité ne signifie pourtant pas qu'elle soit un pur construit. Car elle s'établit d'abord à partir de traces, c'est-à-dire d'empreintes déposées par une action ou un événement. Une trace est toujours trace *de* quelque chose ou *de* quelqu'un ; présence ou représentation présente d'une absence. Paul Ricoeur propose une typologie de la trace et distingue la trace comme empreinte affective, la trace comme empreinte matérielle et la trace comme empreinte documentaire (Ricoeur, 2000). La trace numérique correspond-elle à l'un de ces trois types ? Si nous convenons qu'elle est dématérialisée, au sens où elle n'est pas préhensible comme le sont les objets, et qu'elle n'est pas affective au sens où elle n'est trace dans la mémoire de personne[2], nous pouvons dire qu'elle relève, en revanche, de l'empreinte documentaire. Ricoeur reprend et théorise la définition classique mais fondamentale du document, en tant qu'objet construit, par intention ou par interrogation. Est document tout ce qui peut être interrogé dans l'intention d'y trouver une information sur le passé. En ce sens, une trace numérique est une empreinte documentaire.

Néanmoins, nous comprenons que, tout comme la donnée, elle ne signifie rien par elle-même. Elle ne prend son sens que sous le regard de qui la déchiffre. Déchiffrement par lequel elle devient alors indice, lui-même essentiel à la constitution d'une preuve. Sur ce concept de preuve, la philosophie des sciences nous éclaire tout particulièrement. L'établissement d'une preuve requiert tout ou partie des conditions suivantes (Kaufmann, 1980, p. 249) :

- un énoncé est prouvé s'il est correctement déduit à partir d'autres énoncés ;
- un énoncé est prouvé s'il n'est pas contredit par des faits observés ;
- un énoncé est prouvé s'il est raisonnable de l'accepter.

Cette définition de la preuve est largement inspirée de la pensée d'Aristote. Depuis ses *Analytiques secondes*, on considère en effet qu'un discours n'est admis comme fondé que dans la mesure où il prend la forme d'un développement démonstratif autorisé par la position d'un ensemble de postulats. Mais l'idée aristotélicienne de la démonstration ne met pas seulement en jeu une relation d'inférence entre deux ou plusieurs termes. Elle implique aussi de faire correspondre la hiérarchie des propositions ordonnées par la relation d'inférence à un *ordre réel*. Démontrer, ce n'est pas seulement trouver des conclusions nécessaires selon des schémas d'inférence, mais c'est fonder ontologiquement le choix des postulats. Une preuve est donc toujours problématique en ce qu'elle oscille entre la logique et l'ontologie. Mais elle l'est aussi en ce que ce qui est accepté comme preuve atteste de la disposition morale d'une société donnée à l'égard de l'objet à prouver et de l'autorité de celui qui veut prouver.

Cela étant dit, on peut se demander si une preuve peut être établie uniquement à partir de données. En d'autres termes, le concept de « preuve numérique » a-t-il un sens ? Oui, si l'on considère que les données sont de l'ordre de faits observables. Non, si l'on

considère que la dématérialisation, qui caractérise le numérique, lui retire toute empiricité. Ce sujet de controverse majeure pour le droit actuel fait bien voir comment donnée et preuve sont reliées, et ouvre surtout directement sur le champ de l'éthique. En effet, pourquoi évaluer des données et selon quels critères estimer leur valeur ? Si celle-ci peut s'apprécier par le contenu de la donnée, elle est aussi et surtout fonction du service rendu à son utilisateur. Il est donc indispensable de déterminer quelles données et quelle quantité d'information le concepteur d'un système doit mettre à disposition des utilisateurs. De quelles données ont-ils raisonnablement besoin pour « bien » décider ou « bien » agir ?

Nous espérons avoir montré, par cette clarification conceptuelle, en quoi le continuum trace – donnée – preuve crée un enjeu d'« hyperdocumentarité » (Cormerais, 2018, p. 45) et nous allons appliquer tout notre questionnement à un cas concret.
Dans notre recherche, nous considérons le cas où une organisation prestataire en édition logicielle – Perfect Memory[3] – est mandatée par une organisation, dont la raison sociale est anonymisée, pour apporter une réponse technologique à l'enjeu des *big data* rencontrée par cette dernière. Nous avons pu suivre de l'intérieur les étapes de l'implémentation de la solution – une plateforme de réconciliation des données – et vérifier la promesse du prestataire. Pour la première fois, les silos de l'organisation communiquent entre eux par l'intermédiaire de cette plateforme. En théorie, les données sont intelligibles et la satisfaction client est forte.

2. Problématique

Si nous sortons du strict rapport transactionnel pour nous situer du côté du prestataire qui fournit la solution technologique, un enjeu spécifique apparaît : une technologie capable de réconcilier les données doit-elle le faire pour toutes les données du périmètre confié par le client ? Si la question de savoir si elle le *peut* est tranchée tant du point de vue technique que légal (encore que même ce volet légal soit, en réalité, en pleine structuration), le « doit » fait de la question une interrogation directement éthique. Car le propre des *big data*, c'est de révéler la valeur latente des données, de livrer une réponse sans que la question n'ait été posée à celui qui se retrouve donc à la livrer malgré lui. De sorte que notre problématique pourrait se reformuler ainsi : une technologie capable de réconcilier des données *doit*-elle rendre intelligible toutes les données qu'elle peut légalement traiter ?

Par sa formulation même, cette problématique indique que l'enjeu appréhendé est transverse au droit, aux sciences de l'information et de la communication et à l'éthique. Par « éthique », nous entendons ici une pratique réflexive personnelle et collective sur nos comportements, qui se distingue de la morale (ce qu'une société donnée valide comme étant la « bonne conduite ») et de la déontologie (ensemble de normes édictées par et pour profession).

3. Présentation du terrain

À partir de l'exemple d'une organisation éditrice d'une plateforme de réconciliation de données, nous avons cherché à comprendre le dilemme éthique soulevé par l'accompagnement de l'organisation cliente du point de vue de ses données.

3.1. Parties prenantes

Les parties prenantes de notre étude sont au nombre de deux : une grande entreprise, cliente d'une PME. La raison sociale de la première est anonymisée, celle de la seconde est Perfect Memory, éditeur logiciel dont le siège social est à Lyon 6e. L'organisation cliente, qui produit au quotidien de nombreuses données, exprime au prestataire Perfect Memory un besoin de « réconciliation des données ».

3.2. Présentation de la prestation

À partir du besoin exprimé – en tant que responsable d'une cellule anti-corruption, je souhaite identifier les liens d'intérêt économique existants entre nos salariés et nos fournisseurs (Perfect Memory, 2021a) – le prestataire construit un démonstrateur exploitant plusieurs sources de données. Ce dernier est fondé sur la plateforme logicielle de l'éditeur logiciel, adaptée selon les besoins présents.

Le démonstrateur a pour fonction d'illustrer la capacité du fournisseur Perfect Memory à collecter les différentes sources de données fournies par le client (internes et externes à l'organisation), à les interpréter et à les rendre exploitable par le client, à l'aide d'une interface utilisateur ergonomique. La construction du démonstrateur nécessite plusieurs phases itératives (expression de besoin, spécifications techniques, développement informatique, déploiement en pré-production, recette, déploiement en production, exploitation par le client, support). À l'issue de la prestation, les représentants du client ont accès à une interface utilisateur depuis laquelle lire et traiter les rapports produits par la plateforme à propos de « liens suspicieux » (Perfect Memory, 2021a) identifiés par la réconciliation des données.

3.3. Corpus

De quelles données s'agit-il en réalité ? Le premier cercle de notre corpus est constitué des données réconciliées par le prestataire. La première source est constituée des données fournies par le client et issues de son système d'information « Ressources humaines ». Elle contient la liste des employés et des propriétés relatives à ces

derniers, notamment des données dites personnelles selon le Règlement général sur la protection des données (RGPD). La seconde source fournie par le client consiste en une base de données fournisseurs. Chaque notice est liée à une organisation et contient des informations sur cette dernière, organisées en propriétés. La troisième source interprétée est externe à l'organisation cliente, elle provient d'une entité publique, le tribunal de commerce, qui tient le *Registre du commerce et des sociétés*[4]. La source contient à la fois des fiches, avec des propriétés, mais aussi des documents, tels que des extraits Kbis. La quatrième source interprétée est aussi externe à l'organisation, il s'agit d'une base de données liées et ouvertes (*linked open data*) – Wikidata[5] – éditée par la Wikimedia Foundation. Cette source comporte uniquement des données structurées. Enfin, une dernière source est fournie par le client : la liste des règles métier en lien avec le besoin. Une règle métier est une déclaration de haut niveau structurée qui permet de contraindre, contrôler et influencer un aspect du métier.

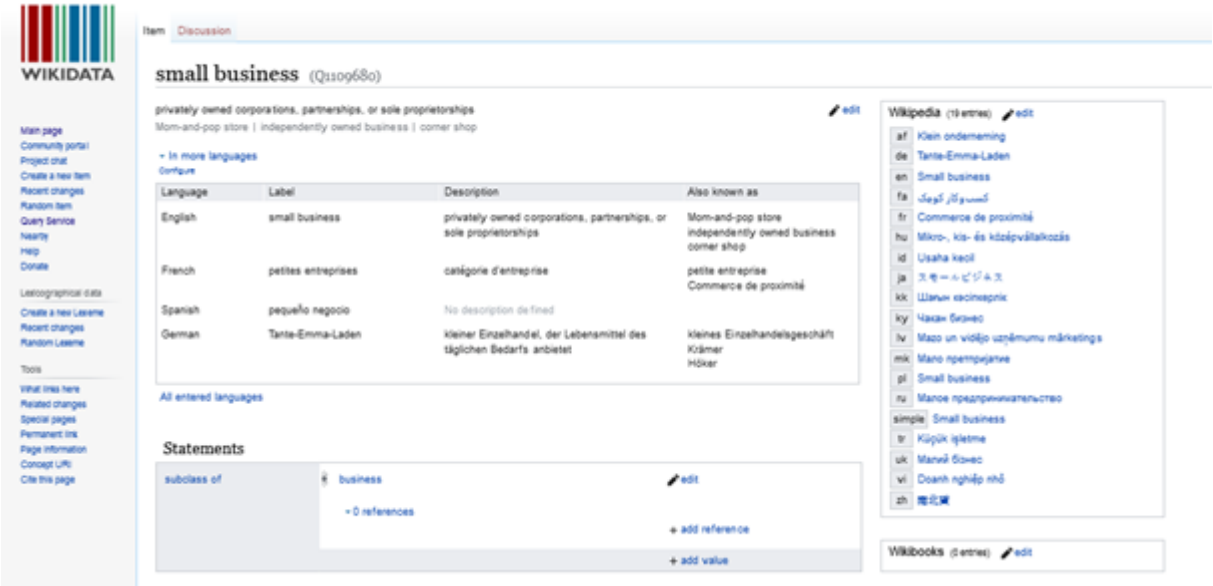


Figure 1 : Parallèlement au portail utilisateur, Wikidata fournit une interface de programmation (API) permettant aux algorithmes d'interroger la base en continu pour rapatrier des notices telles que « Small Business » dans leurs applications ©Wikidata 2021

Le deuxième cercle du corpus est constitué des éléments liés au prestataire Perfect Memory. D'abord, la plateforme de réconciliation des données est un élément à interroger dans ses deux composantes : l'interface utilisateur (*front-end*) du démonstrateur mise à disposition du client puis l'architecture technique (*back-end*) composée notamment du modèle de données, des chaînes de traitement, des services-tiers connectés. Ensuite, au niveau des documents publics produits par le prestataire, nous considérons les éléments de communication du cas d'usage, la déclaration de vie privée, le *Plan d'assurance sécurité*, la méthodologie projet d'intégration du logiciel mais aussi les éléments de communication relatifs à l'éthique et aux valeurs du prestataire. Ces éléments, sous forme de pages web et de documents textes, sont

interrogés au même titre que les données clientes ou que la plateforme du prestataire.

L'ensemble du corpus, articulé autour du cas d'usage « réconciliation des données pour la lutte contre la corruption », est mis gracieusement à disposition des deux auteurs par Perfect Memory, contre signature d'un accord de confidentialité. Ils analysent ainsi les éléments du corpus, à l'aide des approches et concepts mentionnés ci-dessous, dans l'objectif d'apporter des éléments d'approfondissement du dilemme éthique formulé dans la problématique.

4. Méthodologie

Du point de vue de l'ancrage théorique, nous inscrivons notre recherche dans un ensemble d'études dédiées aux « plateformes ». Cette appellation popularisée par l'entrepreneur et essayiste Tim O'Reilly (2011) désigne des formes variées de systèmes techniques et par synecdoque les acteurs économiques commercialisant ces systèmes. Du système d'exploitation à l'application Web en passant par les réseaux sociaux numériques, ces plateformes ont pour point commun une mise en relation entre utilisateurs permise par la médiation d'une infrastructure informatique cristallisant en elle un fort pouvoir sur les usages. La question des plateformes est aujourd'hui abordée par différents prismes, sociologique d'abord à travers l'étude de l'impact des plateformes sur la circulation des biens culturels (Beuscart & Coavoux & Maillard, 2019) ou la transformation des rapports économiques induites par elles (Cardon & Casilli, 2015), communicationnel ensuite à travers un rapprochement entre le procédé d'infomédiation et le phénomène de « plateformisation » (Rebillard & Smyrnaios, 2019). De nombreuses autres études ont été publiées ou sont en cours. Nous avons cherché à savoir si cette notion nous permettait de mieux saisir l'un des éléments clés du corpus – une plateforme de réconciliation des données – et de mettre en perspective sa place médiane dans notre terrain de recherche, entre données et agents des parties prenantes.

Parallèlement, les travaux d'Yves Jeanneret relatifs à une sémiologie des médias et des médiations constituent un cadre fécond pour notre étude. Synthétisant sa démarche, il explique : « je nomme sémiologie l'analyse des formes concrètes et historiques de la médiation et je réserve le terme sémiotique à la conceptualisation de la théorie du signe et de la signification à laquelle recourent ces analyses. Je parlerai donc par exemple de l'analyse sémiologique de Wikipédia et de la sémiotique des écrits d'écran sur laquelle celle-ci se fonde » (Jeanneret, 2014, p.51). Notre cas d'étude comporte ainsi, entre autres, une analyse sémiologique d'un corpus de données, d'un dispositif technique et d'inscriptions fondée sur la sémiotique des écrits d'écran. Trois notions de la sémiologie des médias et des médiations ont retenu notre attention par la densité qu'elles apportent à notre analyse des éléments du corpus.

Issue des études littéraires, la notion de discours d'escorte regroupe les textes qui

accompagnent une œuvre (commentaires, extraits, etc.). Mobilisée par les chercheurs en sciences de l'information et de la communication en l'élargissant à d'autres sources de discours, elle vient aujourd'hui qualifier la dimension performative du discours notamment sur le Web. Le discours d'escorte participe en effet de « *l'imaginaire qui est contenu dans les images, les mots et les phrases, de l'imaginaire qui est porté et travaillé par les discours des acteurs* » (Candel et al., 2019, p.306). Cette notion nous semble féconde dans notre analyse des éléments de communication produits par Perfect Memory.

Interrogeant les modalités d'élaboration de l'écrit contemporain en particulier sur ce nouvel espace qu'est l'écran, Emmanuel Souchier construit en 1996 la notion d'écrit d'écran (Souchier, 1996). L'écrit d'écran, en tant que « *forme particulière que prend l'écrit sur un support numérique équipé de programmes et doté de moyens physiques d'action sur lui (périphériques)* » (Jeanneret, 2014, p.11), s'inscrit dans une théorie générale des cadres : « *l'affichage d'un écrit à l'écran nécessite en effet la mise en abyme d'une série d'au moins quatre cadres successifs* » : le cadre matériel, le cadre système, le cadre logiciel et le cadre document (Jeanneret & Souchier, 2002, p.102). Le constat des auteurs est l'impossibilité de la production d'un texte à l'écran sans outils d'écriture situés en amont. Une mise en abyme du texte est réalisée dans une autre structure textuelle : l'architexte. Du grec *arkhè*, origine et commandement, la notion d'architexte regroupe les « *outils qui permettent l'existence de l'écrit à l'écran et qui, non contents de représenter la structure du texte, en commandent l'exécution et la réalisation* » (Souchier et als, 2003, p.23). La plateforme de réconciliation des données de Perfect Memory est incontestablement un architexte, tandis que ses créateurs, des « maîtres de l'architexte » (Jeanneret & Souchier, 1999, p.106).

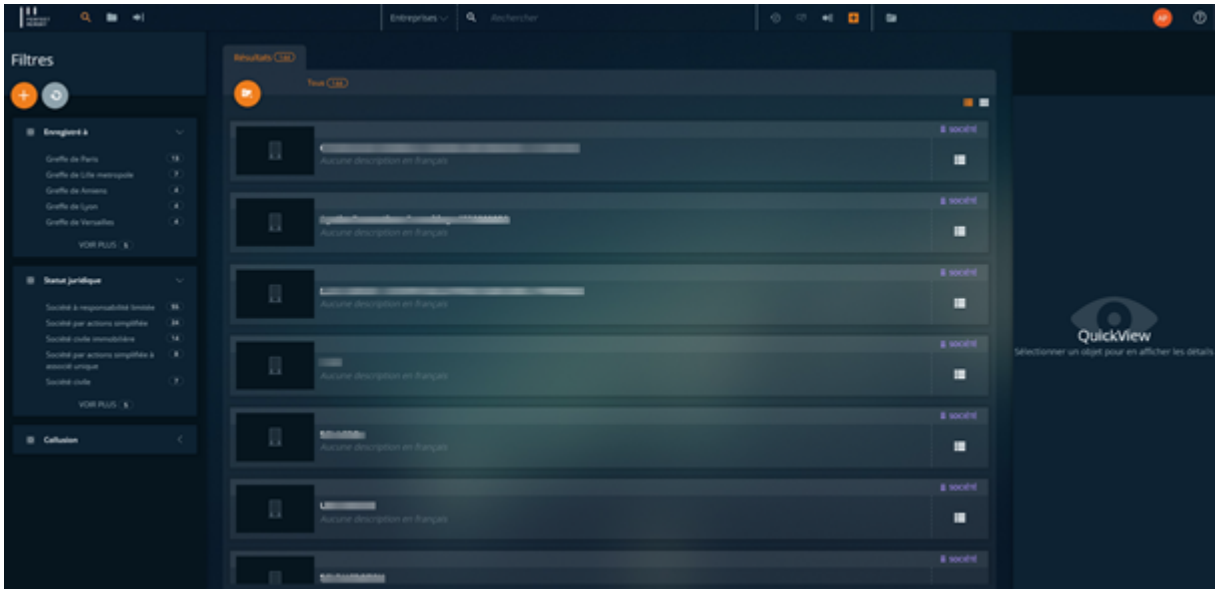


Figure 2 : Textes et signes passeurs coexistent dans cette vue d'accueil d'un des modes de recherche – par « entreprises » – du démonstrateur Anti-corruption ©Perfect Memory 2021

Notre positionnement théorique ne vise pas à interroger les propriétés de la solution technologique, fondée sur la discipline informatique (théorie de conduite de projets informatiques, architecture informatique, ingénierie des connaissances, etc.), mais la façon dont un cas d'usage révèle soit des valeurs embarquées, soit l'absence de référentiel éthique. Il s'agira d'interroger, au fond, la responsabilité d'un prestataire fournisseur d'une technologie capable d'exploiter des traces numériques, en se demandant si le contrat qui le lie à son client et la réglementation en vigueur suffisent à en définir les contours.

Si l'éthique n'est plus un enjeu complètement ignoré des ingénieurs informaticiens depuis une vingtaine d'années (voir les travaux du Software Engineering Ethics Research Institute)[6], la réflexion portant spécifiquement sur une éthique des données est plus récente et encore méconnue en dehors du cercle des spécialistes[7]. Dans le champ des sciences de l'information et de la communication, le questionnement éthique trouve une place dans de nombreux récents travaux dédiés au numérique. L'angle de la 8ème conférence « Document numérique et société » (Balicco & als., 2018) et la création du Groupe éthique et numérique en Information et Communication au sein de la Société Française des Sciences de l'Information et de la Communication (SFSIC), entre autres, appuient le caractère fécond du questionnement éthique dans cette discipline[8].

Pourtant, la notion même d'éthique des données ne semble pas encore avoir reçu de définition bien arrêtée. La proposition la plus complète que nous ayons trouvée, et sur laquelle nous nous appuierons, est faite par les chercheurs Luciano Floridi et Mariarosaria Taddeo :

> « *l'éthique des données peut être définie comme la branche de l'éthique qui étudie et évalue les problèmes moraux liés aux données (y compris la génération, l'enregistrement, la conservation, le traitement, la diffusion, le partage et l'utilisation), aux algorithmes (y compris l'intelligence artificielle, les agents artificiels, l'apprentissage automatique et les robots) et aux pratiques correspondantes (y compris l'innovation responsable, la programmation, le piratage et les codes professionnels), afin de formuler et de soutenir des solutions moralement bonnes (par exemple, les bonnes conduites ou les bonnes valeurs).* » (Floridi & Taddeo, 2016, p.3 – traduction personnelle).

Comment appliquer une telle définition à notre terrain ? C'est ce que les travaux du Think&DoTank danois DataEthics.eu nous aident à faire. En 2018, l'ONG a proposé cinq principes éthiques, associés à un questionnaire d'environ 40 entrées pour les faire vivre et éclairer les dilemmes des organisations concernées. Les principes sont les suivants (Tranberg, 2019, pp.9-12) :

- L'être humain au centre,
- Le contrôle des données personnelles,
- La transparence,

- La responsabilité,
- L’égalité.Résultats

Dans le document *Dataethics*[9], produit par l’ONG, les cinq principes éthiques sont accompagnés d’un questionnaire destiné à évaluer les dilemmes éthiques des organisations. Une fois notre corpus d’éléments constitué, nous avons interrogé ce dernier à partir du prisme des cinq principes et des questions suggérées dans l’ouvrage. La matière obtenue par les réponses aux questions constitue nos résultats préliminaires, que nous discuterons ensuite.

4.1. L’être humain au centre

Ce principe éthique vise à vérifier que les intérêts humains prévalent en permanence par rapport aux intérêts institutionnels et commerciaux. Le postulat défendu est la singularité de l’être humain par rapport aux machines. Ses propriétés – l’unicité, l’empathie, l’autodétermination, l’imprédictibilité, l’intuition et la création – lui garantissent un statut plus élevé que la machine. Si des données sur l’être humain sont manipulées par la machine, les intérêts de ce dernier doivent être au premier plan (Tranberg, 2019, p.9).

L’application de ce principe à notre terrain nécessite de différencier les composantes du traitement des données ayant lieu en amont et pendant la prestation de l’éditeur logiciel.

En amont, nous trouvons les données capitalisées dans le système d’information « Ressources humaines » du client, couvertes par le RGPD. Chaque collaborateur a le droit de regard sur les données personnelles collectées par son employeur. Les données de la base de données « Fournisseurs » n’évoquent que peu l’individu, qui s’efface derrière l’organisation et ses départements. La base du *Registre du commerce et des sociétés*, elle, comporte des notices liées à des personnes et des organisations. Chaque entrée dans la base est conforme à la loi et l’individu ne peut s’y soustraire. Pour finir, les bases de données liées et ouvertes sur le Web, mobilisées par Perfect Memory dans le contexte de l’élévation sémantique d’entités identifiées par les modules d’analyse, comportent en effet des données liées à des personnes – entités « Personne » – néanmoins il s’agit d’informations publiques validées par la communauté de Wikimédiens (contributeurs à l’encyclopédie en ligne Wikipédia, source de la base de données Wikidata). Ainsi, nous pouvons statuer sur le fait que les données sources manipulées par Perfect Memory ne sont pas captées auprès des individus à l’instar de la prédation des données réalisée par les GAFAM, ce sont des données « déclaratives » nécessairement validées par ces personnes lors de l’ajout en base ou ces données se passent de validation car considérées comme appartenant au domaine public.

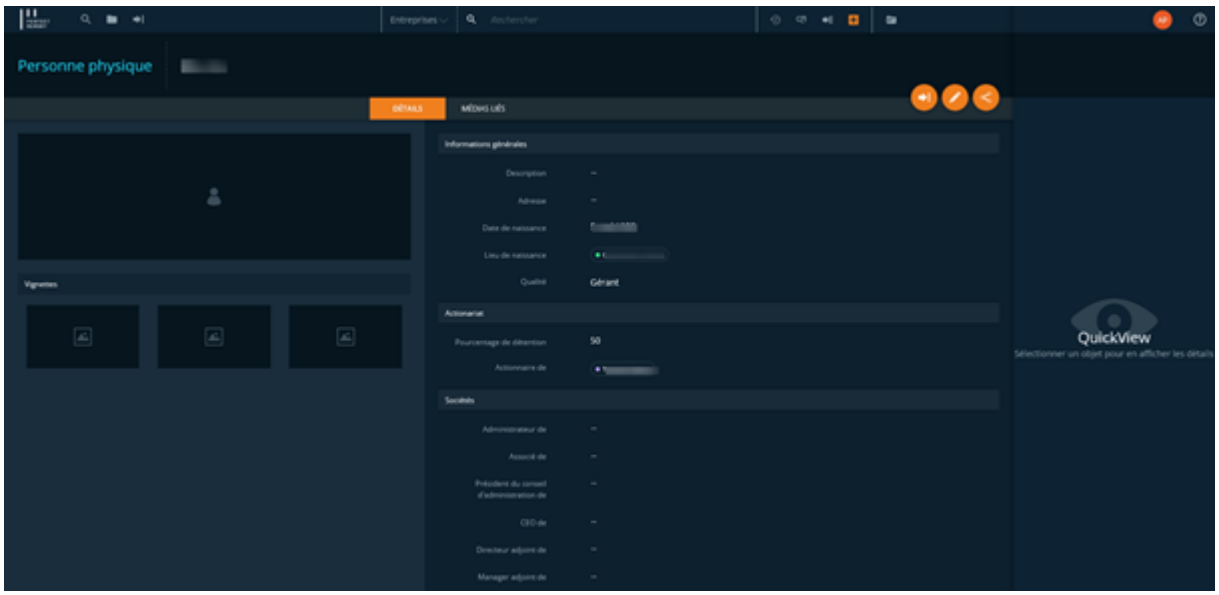


Figure 3 : Une personne physique est dans le démonstrateur une entité à laquelle peut être associée des documents en plus de propriétés spécifiques telles que « Actionnaire de ». ©Perfect Memory 2021

En aval, les utilisateurs du démonstrateur Perfect Memory produisent, eux aussi des données (*logs*) : l'historique de recherche et des actions de création, édition, suppression. Ces données, conservées par le prestataire et mises à disposition du client, ne font pas l'objet de *profiling* par le prestataire. La méthodologie projet inclut, dans la phase d'intégration de la plateforme, une matrice des familles d'utilisateurs et leurs droits associés (Perfect Memory, 2021b). Cette dernière a pour fonction de recenser les besoins en termes d'utilisabilité du service, afin de maximiser la satisfaction cliente. Evolutive, cette dernière est modifiée tout au long du projet à l'initiative du client, non du prestataire.

Concernant la préoccupation relative à la priorisation des intérêts des utilisateurs par rapport aux intérêts commerciaux ou institutionnels, il convient de nuancer les intérêts de l'organisation cliente, des intérêts des utilisateurs du démonstrateur, et enfin des intérêts du fournisseur en tant qu'organisation.

La lutte anti-corruption, en tant que cas d'usage présenté par Perfect Memory sur son site (Perfect Memory, 2021a), est pratiquée par les grandes entreprises à l'échelle d'une équipe ou d'un service. Les intérêts personnels des membres de ce service sont, selon notre position d'observation, difficilement séparable des intérêts économiques et institutionnels de l'organisation qui mandate le service pour cette activité. Il convient néanmoins d'avoir en tête le cas où l'utilisateur du démonstrateur utiliserait ce dernier pour son propre intérêt, il outrepasserait ainsi sa fonction pour son employeur, mais le dispositif technique lui, n'est pas en cause en théorie. Nous pouvons ainsi nous poser la question de la responsabilité de l'éditeur logiciel quant à la mise à disposition d'une plateforme à un utilisateur pouvant l'utiliser en dehors du cadre éthique de son organisation.

4.2. Contrôle des données personnelles

Ce principe éthique vise à ce que l'être humain soit non seulement en contrôle des données qui le concerne mais encapacité par elles. L'auto-détermination d'une personne doit être la priorité des différents traitements de données et la personne doit être impliquée activement vis-à-vis des données produites sur elle. Dans le cas où elles sont traitées, l'être humain doit pouvoir avoir le contrôle principal sur l'usage de ces données.

L'examen du *Plan d'assurance sécurité* (Perfect Memory, 2021d) construit et maintenu par l'éditeur logiciel comporte plusieurs éléments de réponse quant au contrôle des données personnelles.

L'enjeu premier est de garantir que les traitements réalisés sur les données de l'utilisateur ne soient pas réalisés à l'extérieur des terminaux personnels, ou qu'aucun individu ne soit identifiable si ces données sont traitées sur un serveur ou dans le Cloud. L'éditeur garantit le second scénario : les données de l'utilisateur captées par la plateforme sont anonymisées et ne peuvent constituer de traces menant à l'identification d'individus : « *Aucune donnée à caractère personnelle n'est enregistrée sur le système d'information ou les plateformes Perfect Memory* » (Perfect Memory, 2021e). De même, ces données ne sont pas utilisées à des fins de *profiling* ou de prédictions comportementales. L'utilisateur connecté au démonstrateur ne sera pas sujet à une prédation des traces produites. Si les utilisateurs ne voient pas leurs données personnelles contrôlées par la plateforme, qu'en est-il des individus dont les notices apparaissent au fil de la réconciliation des données ? Le principe de rapprocher une notice de personne d'une ou plusieurs notices organisations ne constitue pas en soi du profilage de personnes du fait de l'absence d'une volonté de prédication de comportement. La valeur ajoutée défendue par l'éditeur dans cette démarche de réconciliation des données appliquée à la lutte contre la corruption réside bel et bien dans les « liens cachés entre entités » (Perfect Memory, 2021a) plutôt que dans la construction de profils types.

4.3. Transparence

Ce troisième principe éthique décrit une propriété nécessaire des traitements appliqués aux données : leur transparence pour l'individu. Les objectifs et les intérêts liés aux traitements doivent être compréhensibles par l'individu du point de vue des risques et des conséquences sociales et éthiques de ces traitements.

La transparence est requise à plusieurs niveaux selon ce principe éthique. D'abord, le stockage des données doit être transparent pour l'utilisateur du service de Perfect Memory. Par exemple, les données sont-elles stockées en dehors de l'Union européenne ? L'éditeur se déclare en conformité avec ce principe par son choix d'un hébergeur

français (Perfect Memory, 2021d). Ensuite, si l'éditeur mobilise du *Machine Learning* ou de l'intelligence artificielle, les algorithmes, les critères et les paramètres d'apprentissage doivent pouvoir être explicitées sur demande, notamment des utilisateurs finaux bénéficiant des résultats de l'apprentissage. Sur ce point, la documentation projet fournie par l'éditeur au client comporte des éléments d'intérêt. Il est indiqué qu'au début de tout projet d'intégration du logiciel dans une organisation, les représentants des familles d'utilisateurs finaux sont formés aux principes fondateurs de l'approche dite « sémantique » de l'éditeur Perfect Memory - « Perfect Memory se positionne sur une approche disruptive de l'ingénierie de la connaissance qui mêle IA Heuristique et IA Formelle et matérialisée à travers notre plateforme DAM-as-a-Brain. » (Perfect Memory, 2021c, p. 2). Ainsi, « réconcilier » des données, cela consiste à fournir un référentiel commun (ici une ontologie informatique) qui met à plat les concepts clés (ex. personne, organisation) à la suite d'un rapprochement entre le modèle ancien et celui proposé. Ensuite, il s'agit de paramétrer les algorithmes ayant pour fonction d'interpréter les documents, de produire de l'information à partir de données brutes. D'autres algorithmes ont pour fonction d'élever l'information en connaissances, en rapprochant des informations d'intérêt avec des concepts et entités. Enfin, ces concepts identifiés sont traités par un autre algorithme, venant interroger les bases de connaissances ouvertes pour rapatrier d'éventuelles notices encyclopédiques sur elles. Cette démarche est transparente pour

les représentants des utilisateurs finaux, l'éditeur se positionne ainsi à l'opposé d'un système boite noire et vulgarise le cœur technologique de sa plateforme à destination de ses clients.

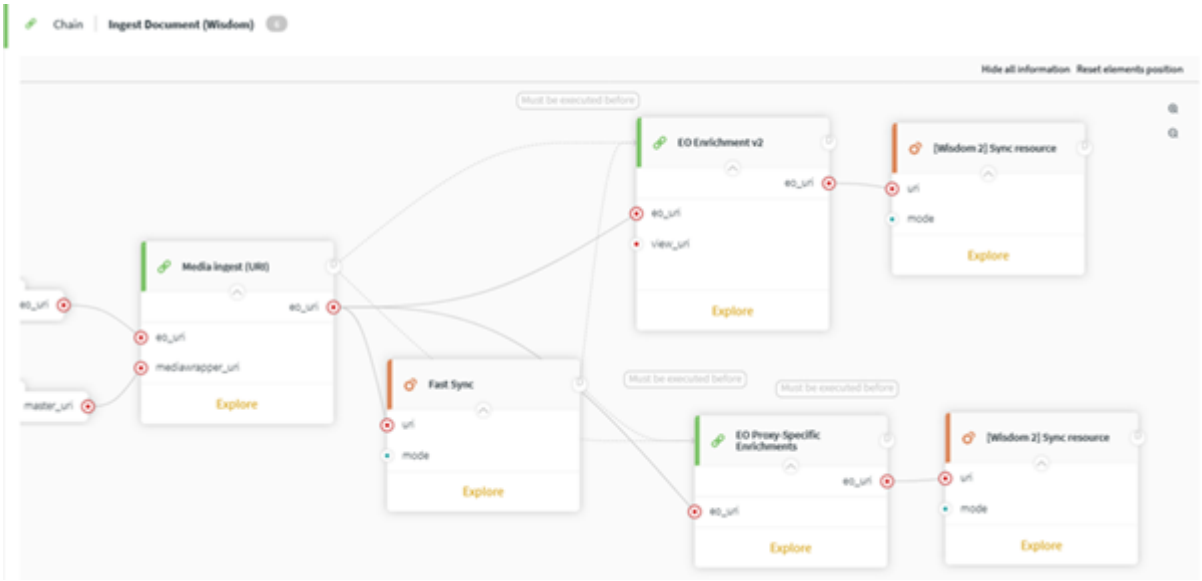


Figure 4 : Derrière ces blocs visuels résident des instructions en code données à la plateforme pour, ici, ajouter un nouveau document dans le démonstrateur. ©Perfect Memory 2021

4.3. Responsabilité

Le principe de responsabilité en éthique des données se décline en plusieurs entrées : anonymat, divulgation nulle de connaissance, vente de données, partage de données, enrichissement des données, ancrage organisationnel et contrôle externe. Selon le référentiel, « *la responsabilité correspond à un usage et une protection des données personnelles réflexifs, raisonnables et systématiques* » (Tranberg, 2019, p. 11). Il précise que non seulement la responsabilité doit faire partie intégrante de chaque composante du traitement que le prestataire, ses fournisseurs et ses partenaires fond des données. Que pouvons-nous déduire du positionnement de l'éditeur logiciel sur le principe de responsabilité à partir des éléments du corpus ?

Sur la question de l'anonymat, nous l'avons vu dans les parties précédentes, l'éditeur est conforme aux différents critères : toute donnée liée à un individu fait l'objet d'une anonymisation dans la chaîne des processus de la réconciliation. De même, les données clientes sont cryptées de bout en bout depuis l'import des sources jusqu'à la restitution dans l'interface utilisateur (Perfect Memory, 2021d).

À propos du principe de divulgation nulle de connaissance, l'éditeur logiciel a fait évoluer ses pratiques d'authentification depuis 2019 en réduisant au strict minimum les informations requises à l'utilisateur pour son inscription au service. Plus récemment, le module d'authentification à chacun des services de l'éditeur logiciel, dont le démonstrateur, s'est vu amélioré par l'utilisation d'un logiciel à authentification unique (SSO) *open source* au travers la gestion par identité et par accès (Perfect Memory, 2021d). La divulgation nulle de connaissance est ainsi bien un principe directeur dans la construction de la plateforme Perfect Memory.

Concernant la vente de données, le positionnement marché de l'éditeur fournissant à des clients une prestation de réconciliation des données ne consiste pas à collecter des données clients sous quelque forme que ce soit dans le but de les vendre à des tierces parties. D'ailleurs, les discours d'escorte véhiculent l'idée que la proposition de valeur de l'éditeur ne réside pas dans la production de nouvelles données pour l'organisation cliente, en d'autres termes apporter au client des données qu'ils n'avaient pas mais, au contraire, faire apparaître ce qui était invisible car noyé dans la masse de données ou simplement inaccessible compte tenu de l'organisation en silos des systèmes d'information, à l'instar du slogan : « Une plateforme cognitive qui révèle l'invisible » (Perfect Memory, 2021a).

De la même manière, les cookies et autres outils de traçage des utilisateurs ne sont pas utilisés par l'éditeur logiciel sur les services proposés, à l'instar du démonstrateur Anti-corruption. Le site vitrine, comme l'indique toute visite sur lui, utilise des cookies selon le respect du RGPD (Perfect Memory, 2021a).

L'enrichissement des données utilisateurs à partir de sources externes, telles que les

réseaux sociaux, des jeux de données achetés ou du *Web scraping* est possible techniquement parlant pour l'éditeur logiciel, en tant que sources additionnelles au périmètre de données à collecter, interpréter et rendre exploitable. Si la possibilité technique n'est pas un frein pour l'utilisation de ces données, l'autorisation des individus concernés par ces données l'est pour l'éditeur logiciel. Les discours d'escorte de l'éditeur mentionnent le cas d'usage d'une organisation « marque-média », produisant et diffusant elle-même de nombreux contenus sur les réseaux sociaux afin de mettre en valeur la marque ou ses produits. L'éditeur a été capable de collecter les documents partagés par la marque et les interactions entre la marque et ses fans. Pour l'utilisation de ces données, se pose au préalable la question de l'autorisation de chaque individu. Toute action d'enrichissement des données mobilisant des données externes doit ainsi se faire en coopération avec les individus concernés par ces données. Dans le cas du démonstrateur Anti-corruption, l'ajout de sources externes réseaux sociaux au périmètre de réconciliation constitue un dilemme éthique pour l'organisation et le prestataire : faut-il demander l'autorisation aux organisations de la base fournisseurs, aux collaborateurs de l'organisation ? Si l'activité de la cellule Anti-corruption est confidentielle, demander une autorisation n'est-il pas contre-productif pour les enquêtes ? En revanche, exploiter ces données externes sans coopération des individus et organisations constitue-t-il une entorse à la loi, notamment au RGPD ? Ce chemin n'ayant pas été emprunté dans le cas du démonstrateur, la question reste en suspens.

La question de la responsabilité se retrouve aussi dans la structuration même de l'organisation prestataire. Si l'éthique est mobilisée en tant que champ conceptuel par la direction de Perfect Memory, notamment dans le référentiel de valeurs (Perfect Memory, 2021a), le management éthique des données n'est mentionné dans aucune fiche de poste du prestataire. Un investissement de cette question par l'équipe Qualité est néanmoins en cours, elle débouchera à terme par une sensibilisation aux enjeux d'éthique des données de chaque département de l'organisation.

Enfin, la manière dont une organisation s'empare de la question de la responsabilité se mesure au degré d'ouverture de l'organisation quant aux audits externes, notamment sur le sujet du traitement des données. Préparant une certification à la norme ISO 27001 (management de la sécurité de l'information), Perfect Memory a inclus au plan d'action un audit externe de l'ensemble du système de management de la sécurité de l'information (SMSI) ainsi construit. À l'issue du processus de mise en conformité, un audit est programmé tous les trois ans (Perfect Memory, 2021f). Par ailleurs, Perfect Memory applique une politique stricte de contrôle de ses fournisseurs et partenaires du point de vue de la sécurité de l'information (Perfect Memory, 2021d). L'entrée éthique pourrait compléter le formulaire de contrôle, dans l'optique d'une application des principes d'éthique des données à l'organisation.

4.4. Égalité

Le dernier principe d'éthique des données mentionné est l'égalité, entendue comme «

un traitement démocratique des données » (Tranberg, 2019, p. 11), c'est-à-dire la conscience que les systèmes de données peuvent maintenir, reproduire ou créer des discriminations. Ce principe suppose de travailler activement à réduire les biais éventuels des algorithmes auto-apprenants.

Dans un premier temps, le statut des plateformes publiques est posé. Les utilisateurs peuvent-ils rentrer en contact avec les représentants du prestataire, y a-t-il des consignes pour l'utilisation de la plateforme, cette dernière est-elle modérée afin d'y supprimer des données personnelles sensibles versées par un utilisateur, l'inscription à la plateforme nécessite-t-il un accord parental pour les utilisateurs mineurs ? Si ces questions ne s'appliquent particulièrement à des plateformes dites *Business to Consumers* (B2C), c'est-à-dire à destination du grand public, plusieurs d'entre elles trouvent des réponses dans le corpus. Le lien entre les utilisateurs et l'équipe Perfect Memory prend forme peut se faire via une autre plateforme dont le nom est anonymisé. Cette dernière, un service d'assistance par tickets, garantit à l'utilisateur que toute demande sera traitée par l'équipe du prestataire, selon le respect du périmètre des conditions de vente. Quant à l'aptitude à l'utilisation du démonstrateur, celle-ci est garantie par la fourniture au client d'un manuel utilisateur (Perfect Memory, 2019). Chacune des fonctionnalités y est décrite, avec des conseils pratiques pour faciliter l'utilisation quotidienne du service.

Dans un second temps, la réutilisation des données peut sembler une question piège pour de nombreux prestataires éditeurs de plateformes. Les données utilisateurs sont-elles ensuite utilisées pour entrainer un algorithme ? Si cette réutilisation des données a lieu, comment garantir qu'elle n'entraîne pas de discrimination ? Le prestataire a-t-il mis en place des moyens de s'assurer que l'utilisation des données n'expose pas les vulnérabilités d'individus ? Comme indiqué plus haut, l'éditeur logiciel Perfect Memory n'utilise pas les données utilisateurs (*logs*). Les algorithmes de l'éditeur sont construits en amont du déploiement des plateformes dont le démonstrateur Anti-corruption, ils mobilisent des règles métiers fournis par le client comme autant d'indications à la machine. Le traitement des données des différentes sources se fait ainsi en lien avec ces règles métiers, et les données produites à l'issue du traitement sont liées à elles. Les algorithmes de Perfect Memory ne sont ainsi pas apprenant, encore moins à partir de données crées lors de session utilisateurs. Le questionnement par rapport aux risques de discrimination ou d'exposition de vulnérabilités d'individus ne n'applique donc pas dans le périmètre d'étude présent.

Enfin, l'accent est placé sur les risques liés à l'utilisation du *Machine Learning* et de l'intelligence artificielle par le prestataire. Celui devrait pouvoir garantir que cette utilisation est au bénéfice de l'individu plutôt que de porter préjudice – de n'importe quelle sorte – à ce dernier. La plateforme Perfect Memory est capable d'utiliser des modules de *Machine Learning* selon des besoins précis de réconciliation de données, à l'instar d'une reconnaissance automatique de visages dans un corpus de vidéos (Perfect Memory, 2021a). Cette démarche possède un prérequis stratégique : pour chaque visage souhaitant être identifié par le module, le prestataire doit détenir au préalable

une base importante de photos au sein desquelles le visage est reconnaissable. Cette base constitue le socle d'entraînement du module qui, ensuite, est capable, en prenant en compte des faux positifs, le visage souhaité au sein de larges corpus d'images fixes ou en mouvement. Le curseur éthique quant à l'utilisation de *Machine Learning* se place, selon notre hypothèse, dans ce prérequis. Le module est incapable de reconnaitre le visage d'un individu sans une solide base iconographique sur ce même individu, il ne peut créer de la connaissance – entité Personne n°X mise à l'image dans le document n°X – qu'à partir d'une base fiable et conséquente. Si cette dernière existe, un dilemme éthique apparaît. Dans les autres cas, vu que le prestataire ne fait appel à des sources de données externes payantes, le prestataire ne peut connecter à la plateforme des modules capables d'identifier des individus souhaités par le client. Aucun préjudice n'est ainsi porté.

5. Bilan

L'éditeur est en conformité avec le référentiel éthique que nous avons choisi de mobiliser. Pourtant, un malaise éthique subsiste, qui nous fait dire que ce référentiel ne permet pas de traiter notre sujet de façon suffisamment problématisante et ce pour trois raisons.

D'abord, dans le cas retenu pour cet article, la demande du client pourrait relever de la règlementation Sapin 2 (JORF, 2016). Nous avons pu observer sur d'autres cas que cette réglementation a trois effets solidaires : elle encourage le recours à l'algorithmie ; ce faisant, elle rend socialement acceptable le principe de traçabilité ; acceptabilité qui renforce elle-même ce que nous pourrions appeler un « paradigme du contrôle », au détriment de la confiance faite à la parole donnée. Comme le dit le philosophe Pierre-Antoine Chardel :

> « *Les technologies qui visent à échanger des informations à notre propos sans que nous ayons à nous exprimer, manifestent un caractère aliénant en signifiant à la place du sujet parlant. D'un point de vue ontologique, la parole est consubstantielle au fait d'être-soi. (...) C'est dans un environnement où nous préservons le devenir des subjectivités, le sens de la parole, du dire et de ses nuances qui font que nous sommes des êtres humains (et non des machines informationnelles), qu'il nous sera possible de préserver un certain niveau de signification dans l'organisation de nos existences* » (Chardel, 2014, pp. 10-11).

Ensuite, le fait que l'éditeur logiciel ne puisse pas garantir que son outil – qui est aujourd'hui un « outil de contrôle » légal – ne se transforme pas en outil d'espionnage. C'est toute l'ambiguïté de la notion de traçabilité, qu'il faudrait continuer d'interroger. D'un point de vue éthique, l'incapacité de l'éditeur à parer au détournement de son outil serait une justification suffisante pour refuser de souscrire à la demande de son client, quand bien même celle-ci est légale[10].

Enfin, les créateurs de la plateforme – maîtres de l'architexte – déclarent souvent, au détour de conversations, qu'ils sont conformes à la règlementation et qu'ils ne font donc rien d'illégal. D'un point de vue éthique, il nous semble que cette posture est démissionnaire. Polarisés par le défi technique, les concepteurs se déresponsabilisent sur le plan éthique.

Fort de ce constat, nous souhaitons ouvrir sur deux recommandations.

En matière de collecte des données, certains auteurs parlent déjà aujourd'hui de « consentement libre et éclairé » des utilisateurs (Peraya, 2019). Dans l'esprit de ce principe, l'entreprise pourrait pousser la transparence jusqu'à informer ses collaborateurs de l'utilisation qui est faite de leurs données via la plateforme de réconciliation de données.

Par ailleurs, nous suivons les propositions de Célia Zolynski, professeur de droit privé, concernant des modes de régulation *ex ante* :

> « *[les opérateurs] devraient être tenus d'un devoir de rendre des comptes imposant de décrire les garanties qu'ils mettent en œuvre pour assurer le respect des droits des tiers selon une logique d'accountability. (…) Il convient, en même temps, de porter la reconnaissance de nouvelles obligations à mettre à la charge de ces opérateurs à l'image des obligations de transparence et de loyauté préconisées par le Conseil national du numérique et le Conseil d'état et, plus spécifiquement, d'imposer un contrôle de l'usage des algorithmes, par exemple par un tiers certificateur, un « algorithmiste », sur le modèle des commissaires aux comptes* » (Zolynski, 2015, p. 26). Cette recommandation fait écho à la notion de commissariat aux données, proposée par l'avocat Matthieu Bourgeois en 2017 (Bourgeois, 2017).

En conclusion, cette première étude de cas nous indique que l'éthique des données est un enjeu interdisciplinaire, à la croisée de l'ingénierie informatique, du droit et de la philosophie. Un défi propre à être relevé par les responsables qualité des éditeurs logiciels ? Peut-être. À moins qu'un nouveau métier dédié ne soit appelé à apparaître.

6. Bibliographie

Anderson, C. (2008). *The End of Theory: The Data Deluge Makes The Scientific Method Obsolete*, Wired. https://www.wired.com/2008/06/pb-theory/

Béranger, J. (2014). *Big Data et données personnelles : vers une gouvernance éthique des algorithmes*, Paris Tech Review, Paris. https://www.paristechreview.com/

Beuscart, J.-S., Coavoux, S., Maillard, S. (2019). Les algorithmes de recommandation musicale et l'autonomie de l'auditeur. Analyse des écoutes d'un panel d'utilisateurs de streaming. *Réseaux*, n° 213, 17-47.

Bourgeois, M., & Bounedjoum, A. (2017). *Droit de la donnée: principes théoriques et approche pratique*. LexisNexis.
Candel, E., Gomez-Mejia, G., Jeanne-Perrier, V., & Souchier, E. (2019). *Le numérique comme écriture*. Armand Colin. https://hal.archives-ouvertes.fr/hal-01982714
Cardon, D., Casilli, A. (2015). *Qu'est-ce que le Digital Labor ?* Bry-surMarne, INA.
Chardel, P.-A. (2014). Données personnelles et devenir des subjectivités. Questions d'éthique. *Sécurité et stratégie*, 2014/2 vol. 17, 5-11.
Cormerais, F. (2018). Éthique de l'intention et gouvernement des données. Dans : Laurence Balicco éd., *L'éthique en contexte info-communicationnel numérique: Déontologie, régulation, algorithme, espace public* (pp. 39-50). Louvain-la-Neuve, Belgique: De Boeck Supérieur. https://doi.org/10.3917/dbu.balic.2018.01.0039"
Dictionnaire de l'Académie Française. 1: A - Enz (9. éd). (2005). Impr. Nationale.
Floridi, L., & Taddeo, M. (2016). What is data ethics? *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, *374*(2083), 20160360. https://doi.org/10.1098/rsta.2016.0360
Gotterbarn, D. (2002). Software Engineering Ethics. Dans Marciniak. J.J., *Encyclopedia of Software Engineering*, Wiley Online Library. DOI:10.1002/0471028959.sof314.
Guchet, X. (2013) Technique et philosophie. Une contribution de François Russo, *Bulletin de la Sabix*, 53-2013, 7-15. DOI:10.4000/sabix.1055
Hansson, S. (2002). Les incertitudes de la société du savoir. *Revue internationale des sciences sociales*, 171, 43-51. https://doi.org/10.3917/riss.171.0043
Jeanneret, Y. (2014). *Critique de la trivialité: les médiations de la communication, enjeu de pouvoir*. Éditions Non standard.
Jeanneret, Y., & Souchier, E. (1999). Pour une poétique de « l'écrit d'écran ». *Xoana*, 6, 97-107
Jeanneret, Y., & Souchier, E. (2001). Présentation. *Communication & Langages*, *128*(1), 33-33. https://doi.org/10.3406/colan.2001.3072
Jeanneret, Y., & Souchier, E. (2002). Écriture numérique ou médias informatisés ?. *Dossier pour la science, hors-série n°33 octobre-janvier*, 100-105
Journal officiel de la République française (2016), LOI no 2016-1691 du 9 décembre 2016 relative à la transparence, à la lutte contre la corruption et à la modernisation de la vie économique, *JORF n°0287 du 10 décembre 2016*, https://www.legifrance.gouv.fr/eli/jo/2016/12/10
Kaufmann, N. (1980). Les trois niveaux de la preuve dans les sciences empirico-formelles. *Philosophiques*, *7*(2), 247-265. https://doi.org/10.7202/203142ar
Lurie, Y., Mark. S. (2015). Professional Ethics of Software Engineers: An Ethical Framework. *Science and Engineering Ethics*, vol. 22-2, June 2015. http://dx.doi.org/10.1007/s11948-015-9665-x
Merzeau, L. (2009). Du signe à la trace : l'information sur mesure. *Hermès, La Revue-Cognition, communication, politique*, CNRS-Editions, 23-29. halshs-00483292f
O'Reilly, T. (2011). Government as a Platform. *Innovations: Technology, Governance, Globalization*, vol. 6, n° 1, 13-40.
Perfect Memory (2019). Manuel utilisateur Trouvé™
Perfect Memory (2021a). Anti-bribery Anti-corruption. https://www.perfect-memory.com/our-solutions/anti-bribery-anti-corruption/

Perfect Memory (2021b). Méthodologie projet (confidentiel)
Perfect Memory (2021c). Perfect Academy. Plan de formations certifiantes (confidentiel)
Perfect Memory (2021d). Plan d'assurance sécurité (confidentiel)
Perfect Memory (2021e). Réponses à propos du RGPD (confidentiel)
Perfect Memory (2021f). Système de management de la sécurité de l'information (confidentiel)
Rebillard, F., Smyrnaios, N. (2019). Quelle "plateformisation" de l'information ? Collusion socioéconomique et dilution éditoriale entre les entreprises médiatiques et les infomédiaires de l'Internet, *tic&société*, vol. 13, n° 1-2, 247-293.
Ricœur, P. (2003). *La mémoire, l'histoire, l'oubli*. Ed. du Seuil.
Software Engineering Ethics Research Institute. (1999). *Code d'éthique et déontologique de l'ingénieur logiciel*. IEEE & ACM. https://studylibfr.com/doc/2899636/code-d-%C3%A9thique-et-d%C3%A9ontologique-de-l-ing%C3%A9nieur-logiciel
Souchier, E. (1996). L'écrit d'écran, pratiques d'écriture & informatique. *Communication & Langages*, *107*(1), 105-119. https://doi.org/10.3406/colan.1996.2662
Souchier, E., Jeanneret, Y., & Marec, J. L. (2003). *Lire, écrire, récrireâ¯: objets, signes et pratiques des médias informatisés*. BPI. https://halshs.archives-ouvertes.fr/halshs-00425754
Tranberg, P. (2019). *Dataethics: principles and guidelines for companies, authorities & organisations*. PubliShare]â¯: with Spintype.com.
Zolynski, C. (2015). Big Data : pour une éthique des données. *A.D.B.S. | « I2D - Information, données & documents »*, 2015/2, vol. 52, 25-26.

7. Notes

[1] N'étant pas juristes, nous mentionnerons ce qu'il en est du droit à titre descriptif et non analytique.
[2] Ces deux suppositions pourraient être philosophiquement discutées mais nous choisissons délibérément ici de concentrer notre propos sur ce qui relie la trace à la donnée et à la preuve.
[3] https://www.perfect-memory.com/
[4] https://www.infogreffe.fr/
[5] https://www.wikidata.org/wiki/Wikidata:Main_Page
[6] Lors de la première publication de l'Encyclopédie de l'ingénierie logicielle, l'un des contributeurs, rappelle que « dans toute société donnée, l'éthique régit le comportement individuel dans les circonstances où il ne peut être ou n'a pas encore été réglementé par la loi. » (Gotterbarn, 2002 – traduction personnelle)
[7] Notre recherche nous a montré que très peu d'articles en langue française existent sur le sujet.
[8] Nous renvoyons le lecteur au stimulant appel à communication lancé par le groupe plus tôt cette année : https://www.sfsic.org/aac-publication/questionner-lethique-depuis-les-sic-en-contexte-nu

merique/

[9] https://dataethics.eu/data-ethics-principles/

[10] Toute proportion gardée, nous pourrions faire le parallèle avec le cas du scientifique Joseph Rotblat, connu pour avoir quitté le projet Manhattan avant même le bombardement d'Hiroshima, pour des raisons morales.