Comment (ré)concilier RGPD et big data ?

Delforge, Antoine

Published in:
Revue du Droit des Technologies de l'information

Publication date:
2018

Document Version
le PDF de l'éditeur

Comment (ré)concilier RGPD et big data ?

Antoine Delforge[1]

Depuis plusieurs années maintenant, de nombreuses entreprises souhaitent se lancer dans des stratégies de big data afin de valoriser au maximum les données qu'elles possèdent. Bien souvent, parmi ces données se trouvent des données d'un genre un peu particulier, des données à caractère personnel. Il faut alors concilier RGPD et big data, deux termes que tout semble opposer. L'exercice est certes délicat mais est-il pour autant insurmontable ?

INTRODUCTION

Ce n'est un secret pour personne, le volume de données créées chaque jour a fortement augmenté ces dernières années. Les techniques et les capacités d'analyse ont elles aussi énormément progressé et il est devenu possible de tirer de nombreuses informations à partir de données brutes, plus ou moins structurées à l'avance, et *a priori* sans valeur particulière.

Afin de refléter cette évolution significative, est apparu le terme big data, terme dont la définition varie parfois légèrement. Il signifie en effet tantôt les données elles-mêmes, tantôt le type particulier d'opération effectuée à partir de celles-ci ou encore le phénomène global d'accroissement de la quantité de données qui sont créées et utilisées[2].

Aujourd'hui, énormément d'organisations (tant les entreprises privées que les États) utilisent le big data pour rendre plus efficaces leurs procédures, pour mieux connaître leurs clients, pour tenter de prédire certains comportements…

Le big data est donc apparu comme une opportunité inattendue de valoriser la masse parfois insoupçonnée de données contenues dans leurs serveurs. Bien souvent, en effet, les données existaient et étaient stockées pour d'autres raisons, voire même parfois sans raison. Beaucoup d'organismes ont dès lors été tentés de « faire du big data »[3] sur leurs jeux de données.

Cependant, se lancer dans ce type de processus ne peut se concevoir sans tenir compte du cadre légal applicable aux traitements de données, et particulièrement du règlement général sur la protection des données[4] (ci-après « RGPD »). De fait, ce règlement européen vise à encadrer les traitements de données à caractère personnel et régulièrement les données exploitées dans le cadre d'analyse big data sont des données

1 Chercheur au Centre de Recherche Information, Droit et Société (Crids), Université de Namur.

2 Datatilsynet (L'autorité de protection des données norvégienne), *Big data – Privacy principles under pressure*, september 2013, disponible sur https://www.datatilsynet.no/globalassets/global/english/big-data-engelsk-web.pdf.

3 Expression volontairement floue englobant assez largement les activités d'analyse de données massives.

4 Règlement (UE) 2016/679 du Parlement européen et du Conseil du 27 avril 2016 relatif à la protection des personnes physiques à l'égard du traitement des données à caractère personnel et à la libre circulation de ces données, et abrogeant la directive 95/46/CE (règlement général sur la protection des données), *J.O.*, L 119, 4 mai 2016 ; consolidé suite au rectificatif du 23 mai 2018.

qualifiées par le règlement comme étant « à caractère personnel ».

Dans le cadre de cette contribution, nous nous attacherons d'abord à préciser le champ d'application du RGPD[5] afin de déterminer les cas dans lesquels il est nécessaire de prendre en compte les prescrits de ce nouveau texte européen. Nous passerons en revue différents points du RGPD qui paraissent *a priori* difficiles à respecter quand on « fait du big data »[6].

I. LE RGPD, UN TEXTE DEVENU QUASI INCONTOURNABLE

Avant d'analyser les règles et principes contenus dans le RGPD, il est nécessaire d'étudier son champ d'application, qui est beaucoup plus large qu'on ne pourrait le penser au premier abord.

A. Le champ d'application matériel

Le RGPD s'applique à tout « traitement de données à caractère personnel, automatisé en tout ou en partie […][7] ».

Avant d'analyser plus en détail la notion de *traitement* et de *donnée à caractère personnel*, précisons d'emblée que le terme *automatisé* fait référence aux technologies informatiques permettant de traiter des données.

Le RGPD se voulant neutre technologiquement[8], il s'applique à toutes les technologies actuelles (les techniques de big data y compris) et à venir permettant « d'accéder à une ou plusieurs données enregistrées dans un vaste ensemble […], les sélectionner, les extraire, les associer, les modifier, etc. sans qu'il soit nécessaire que les données aient fait l'objet d'une structuration préalable pour arriver à ces résultats »[9].

Le RGPD vise à imposer un cadre légal aux traitements[10] *de données à caractère personnel*. Le

5 Nous ne traiterons pas des exceptions prévues à l'article 2, § 2, du RGPD. Le champ d'application territorial ne sera quant à lui qu'évoqué dans un souci d'exhaustivité, mais ne sera pas analysé en profondeur. Sur ces différents points, nous renvoyons à la contribution de C. de Terwangne, « Définitions clés et champ d'application du RGPD », in *Le règlement général sur la protection des données : Analyse approfondie*, Bruxelles, Larcier, à paraître.

6 Nous ne parlerons pas ici des obligations du responsable du traitement. Ceci dépasse le cadre de cette contribution et nécessiterait une étude trop détaillée de chacune d'entre elles. Nous pouvons néanmoins attirer l'attention sur le fait que, par nature, effectuer des analyses big data signifie traiter un volume de données conséquent. Ceci aura pour conséquence que ce type de traitement risque probablement d'obliger le responsable de ce traitement à nommer un délégué à la protection des données (voy. art. 37 du RGPD), à réaliser des analyses d'impact préalablement à chaque nouveau traitement (voy. art. 35 du RGPD), à adopter un niveau de sécurité élevé (voy. art. 32 du RGPD)… Pour une étude complète de ces obligations, nous renvoyons à l'ouvrage *Le règlement général sur la protection des données : Analyse approfondie*, *op. cit.*, à paraître.

7 […] ainsi qu'au traitement non automatisé de données à caractère personnel contenues ou appelées à figurer dans un fichier. Le RGPD s'applique aussi dans certains cas aux traitements de données à caractère personnel qui s'effectuent sur papier si les données sont contenues ou sont appelées à figurer dans un *fichier*. Nous ne nous attarderons pas sur cette notion de *fichier* vu que les traitements de données de type big data se font par nature de manière numérique. Pour plus de détails sur la notion de *fichier,* voy. notamment C. de Terwangne, « Définitions clés et champ d'application du RGPD », *op. cit.*, §§ 8 et s.

8 Cons. 15 du RGPD.

9 C. de Terwangne, « Définitions clés et champ d'application du RGPD », *op. cit.*, § 7.

10 Par *traitement*, on entend « toute opération ou tout ensemble d'opérations effectuées ou non à l'aide de procédés automatisés et appliquées à des données ou des ensembles de données à caractère personnel, telles que la collecte, l'enregistrement, l'organisation, la structuration, la conservation, l'adaptation ou la modification, l'extraction, la consultation, l'utilisation, la communication par transmission, la diffusion ou toute autre forme de mise à disposition, le rapprochement ou l'interconnexion, la limitation, l'effacement ou la destruction », art. 4, 2), du RGPD. La définition

règlement définit la notion de *donnée à caractère personnel* comme « toute information se rapportant à une personne physique identifiée ou identifiable […] »[11].

Cette définition volontairement large vise donc *toute information* – quelles que soient sa nature ou sa forme[12] et la manière dont celle-ci est organisée/structurée[13] ou présentée, pour autant qu'elle soit relative à une personne physique identifiée ou identifiable.

L'information doit se rapporter à une *personne physique* de sorte que les traitements de données sur des personnes morales[14] ou des personnes décédées échappent au cadre réglementaire imposé par le RGPD.

Cette information ne doit pas forcément concerner la vie privée d'une personne. Le caractère confidentiel ou non de l'information n'est pas pertinent[15]. Les informations relatives à la vie professionnelle d'un individu[16] et les informations publiques[17] sont, à titre d'exemple, considérées comme des données à caractère personnel au regard du RGPD.

De plus, l'information doit être *relative à une personne identifiée ou identifiable*, mais ne doit pas forcément concerner directement cette dernière. L'information peut notamment se rapporter en premier lieu à un objet, un animal… mais être indirectement relative à une personne. Tel sera notamment le cas d'une plaque d'immatriculation qui est une information sur une voiture, mais constitue également une information sur le propriétaire de cette voiture[18]. Les analyses big data sur des données récoltées à partir d'objets connectés peuvent donc tomber dans le champ d'application matériel du RGPD si ces dispositifs sont installés sur une personne.

Enfin, dernier élément de la définition, il est nécessaire que cette information soit relative à une *personne identifiée ou identifiable*, à savoir une personne pouvant « être identifiée, directement ou indirectement, notamment par référence à un identifiant, tel qu'un nom, un numéro d'identification, des données de localisation, un identifiant en ligne, ou à un ou plusieurs éléments spécifiques propres à son identité physique, physiologique, génétique, psychique, économique, culturelle ou sociale »[19]. On peut également citer comme identifiant le numéro de registre national[20], une adresse IP ou des cookies…

Conceptuellement, cela signifie que la personne peut être individualisée, ciblée, isolée des autres, reconnues. Il ne s'avère donc pas nécessaire de savoir qui est cette personne

étant à ce point large qu'elle englobe toutes les opérations possibles sur des données.

11 Art. 4, 1), du RGPD. Pour une analyse plus complète de cette notion, voy. notamment Groupe 29, *Avis 4/2007 du 20 juin 2007 sur le concept de données à caractère personnel*, WP 136.

12 Texte, image, son, chiffre, film…

13 Qu'importe que les données soient structurées (comme c'est le cas dans une base de données) ou non (image, texte…).

14 Notons que bien souvent, les personnes morales sont incarnées par des personnes physiques. Il faudra donc être attentif dans ce genre de cas de figure.

15 Le nom d'une personne constitue déjà une donnée à caractère personnel. Tel est également le cas de son adresse, de sa profession et de son numéro de téléphone.

16 Le numéro de téléphone professionnel de la personne ou sa fonction seront qualifiés de données à caractère personnel.

17 Les informations de notoriété publique ou rendues publiques par la personne elle-même ou un tiers. Les données librement accessibles sur internet ne sont, par exemple, pas exclues du champ d'application du RGPD.

18 C. DE TERWANGNE, « Définitions clés et champ d'application du RGPD », *op. cit.*, § 2 et références citées.

19 Art. 4, 1), du RGPD.

20 Ce numéro ne peut être utilisé que conformément à la loi du 8 août 1983 organisant un registre national des personnes physiques.

(connaître son nom). Il suffit qu'elle puisse être « traitée différemment des autres »[21].

Si ce n'est pas le cas et qu'il est matériellement impossible – ou au prix d'efforts disproportionnés compte tenu de la nature des données – de procéder à cette individualisation, alors les données sont considérées comme *anonymes* et ne sont alors plus protégées par le RGPD.

Pour apprécier le caractère d'identifiabilité de la personne, « il convient de prendre en considération l'ensemble des moyens raisonnablement susceptibles d'être utilisés par le responsable du traitement ou par toute autre personne pour identifier la personne physique directement ou indirectement »[22].

Cette appréciation reste une question éminemment factuelle qu'il reviendra à chaque responsable du traitement[23] d'effectuer pour chaque donnée qui pourrait être qualifiée de *donnée à caractère personnel*. Il est important de rappeler que derrière un numéro peut parfois se cacher un individu, une personne identifiable pour reprendre la terminologie du RGPD, et ce n'est qu'en vérifiant à quoi correspond chaque donnée utilisée qu'il est possible d'arriver à la conclusion que le traitement de données envisagé échappe au RGPD.

Cette analyse méthodique est d'autant plus cruciale que les technologies d'analyse big data permettent de déduire des informations, des corrélations, à partir de données dont le *data scientist* (spécialiste en analyse de données) n'a pas forcément besoin de savoir précisément à quoi elles correspondent. Les données sont en effet parfois vues comme de simples variables d'une équation dont il n'est pas nécessaire de connaître en détail la nature. Il faut donc veiller à toujours vérifier ce à quoi ou à qui se rapporte chaque donnée utilisée[24]. Il est particulièrement risqué de procéder à des analyses de données sans savoir à quoi chaque variable correspond. À défaut, il existera toujours un risque de ne pas détecter la présence de données à caractère personnel. Cette négligence équivaudra inévitablement à violer certains principes et/ou obligations contenus dans le RGPD et donc à s'exposer à des sanctions potentiellement très lourdes[25].

Ce phénomène va de pair avec une autre tendance apparue suite à la démocratisation et au perfectionnement des technologies de l'information, et notamment des technologies big data. Il devient de plus en plus rare que les données soient considérées comme anonymes

21 En anglais, le terme employé est « single out ». Voy. C. DE TERWANGNE, « Définitions clés et champ d'application du RGPD », *op. cit.*, § 5.

22 Cons. 26 du RGPD.

23 Le responsable du traitement est « *la personne physique ou morale*, l'autorité publique, le service ou un autre organisme *qui, seul ou conjointement avec d'autres, détermine les finalités et les moyens du traitement* ; lorsque les finalités et les moyens de ce traitement sont déterminés par le droit de l'Union ou le droit d'un État membre, le responsable du traitement peut être désigné ou les critères spécifiques applicables à sa désignation peuvent être prévus par le droit de l'Union ou par le droit d'un État membre », art. 4, 7), du RGPD » (mis en italique par nos soins). Sur cette notion, voy. notre contribution, « Les obligations générales du responsable du traitement et la place du sous-traitant », in *Le règlement général sur la protection des données : Analyse approfondie*, *op. cit.*, à paraître.

24 Il est également nécessaire d'examiner si les données ne tombent pas dans des catégories particulières de données (données relatives à la santé, l'orientation sexuelle ou politique…), lesquelles sont soumises à des régimes spécifiques plus restrictifs et contraignants. Voy. notamment art. 9 et 10 du RGPD.

25 Les amendes administratives, pour le secteur privé, peuvent se chiffrer jusqu'à 20.000.000 d'euros ou, jusqu'à 4 % du chiffre d'affaires annuel mondial, voy. art. 83, § 5, du RGPD. Le secteur public (les autorités publiques pour reprendre la définition contenue à l'article 5 de la loi belge du 30 juillet 2018 relative à la protection des personnes physiques à l'égard des traitements de données à caractère personnel) n'est pas soumis aux amendes administratives, voy. art. 221, § 2, de ladite loi.

et bien souvent ces données *a priori* anonymes sont des données à caractère personnel. Il est en effet devenu dans certains cas relativement facile de réidentifier quelqu'un à partir de données *a priori* anonymes grâce à d'autres données. De fait, grâce à la multiplication des sources d'information accessibles et au perfectionnement des technologies de croisement et d'analyse de données, il est parfois possible de réidentifier ou d'isoler un individu, en recoupant des données qui, si elles sont prises séparément, sont considérées comme des données *anonymes*. Aucune des données à elle seule ne permet directement une réidentification, mais en les croisant, il s'avère possible de cibler un individu en particulier. Cette évolution des techniques rend donc plus compliqué d'anonymiser des données à caractère personnel[26].

Les techniques d'anonymisation ont pour principe de supprimer les éléments directement identifiants et quasi identifiants (permettant en les recoupant avec d'autres données de réidentifier une personne)[27]. Pour ce faire, les données sont soit légèrement modifiées ou alors généralisées (on ne donne plus l'âge d'une personne, mais une tranche d'âge…). Ces techniques altèrent donc la qualité des données et parfois il est impossible d'anonymiser certaines données sans leur faire perdre tout leur intérêt.

De plus, cette évaluation du risque de réidentification doit se faire préalablement à la mise en place d'un nouveau traitement de données et tout au long de celui-ci. En effet, il se peut que des données considérées initialement comme anonymes[28] se transforment en données à caractère personnel suite à l'apparition de nouvelles techniques ou l'accès à de nouveaux jeux de données. Certaines données anonymes ou anonymisées peuvent donc, du jour au lendemain, (re)devenir des données à caractère personnel, ce qui imposera alors de respecter à ce moment-là le RGPD.

Dès lors, pour tout traitement de données où il est envisagé de croiser des données, d'analyser des volumes de données conséquents, surtout si le responsable du traitement a à sa disposition des outils de big data, ce qui amplifie le risque de réidentification, il faut s'assurer d'avoir correctement qualifié les données exploitées et ne pas hésiter à vérifier régulièrement que cette qualification ne doit pas évoluer et reste d'actualité. Le RGPD ne verra pas à s'appliquer à chaque fois, mais il est imprudent de l'exclure par défaut dans certains domaines d'activités où *a priori* il n'est pas question de données à caractère personnel.

B. Le champ d'application territorial

Conformément à son article 3, le RGPD s'applique « au traitement des données à caractère personnel effectué dans le cadre des activités d'un établissement d'un responsable du traitement ou d'un sous-traitant sur le territoire de l'Union, que le traitement ait lieu ou non dans l'Union ». Cela signifie que le règlement s'appliquera, qu'importe la nationalité des personnes dont les données sont traitées. Le fait que le jeu de données utilisé ne concerne pas des citoyens européens ou des personnes se trouvant sur le territoire de l'Europe, mais

[26] Précisons que le RGPD, au nom du principe de minimisation, impose d'anonymiser les données à caractère personnel utilisées si cela est possible, art. 5, c), du RGPD. Voy. *infra*, II, B.

[27] Pour plus d'informations à ce sujet, nous renvoyons notamment aux travaux du Groupe 29, *Avis 05/2014 du 10 avril 2014 sur les techniques d'anonymisation*, WP 216. Pour une étude plus approfondie et technique de ces mesures dans un contexte big data, voy. le rapport de l'ENISA, *Privacy by design in big data : An overview of privacy enhancing technologies in the era of big data analytics*, december 2015, disponible sur www.enisa.europa.eu.

[28] Données pour lesquelles le lien entre l'information et la personne concernée par cette information est complètement et irrémédiablement rompu.

des citoyens américains, par exemple, n'est pas un critère pertinent pour le champ d'application territorial du RGPD. Dès que le traitement de données est effectué par un responsable du traitement établi au sein de l'Union européenne, les règles contenues dans le RGPD doivent être respectées, qu'importe où les données sont traitées.

Le RGPD s'appliquera également, peu importe le lieu d'établissement du responsable du traitement, si ce dernier traite des « données de personnes concernées qui sont dans l'Union européenne lorsque soit les activités de traitement concernent l'offre de biens ou services à cette personne soit les activités de traitement concernent la surveillance des comportements de ces personnes concernées pour autant que ce comportement a lieu au sein de l'Union européenne »[29] [30].

II. LES DIFFICULTÉS PARTICULIÈRES POSÉES PAR LES TRAITEMENTS DE DONNÉES DE TYPE BIG DATA

La législation relative à la protection des données a récemment connu, avec l'entrée en application du RGPD, une mise à jour des plus conséquentes afin de mieux s'adapter aux réalités du monde d'aujourd'hui, monde numérique où les données sont devenues pour de nombreuses entreprises et même pour les États une matière première qui doit être exploitée au maximum.

Ce nouveau règlement est certes plus adapté au monde actuel et aux techniques utilisées à présent dans l'exploitation des données, mais il n'empêche que la philosophie du RGPD et les grands principes contenus dans celui-ci peuvent sembler parfois en contradiction avec la logique même de ce qu'on appelle le big data.

Dans ce chapitre nous passerons donc en revue différents éléments[31] qui peuvent s'avérer potentiellement problématiques quand un organisme, public ou privé, souhaite exploiter des volumes importants de données (à caractère personnel), grâce aux nouvelles techniques d'analyse big data. Certains vont jusqu'à dire que le RGPD et le big data sont incompatibles[32], d'autres préfèrent parler de *challenge* pour la législation européenne relative à la protection des données[33].

Il est vrai que le RGPD peut constituer un frein dans certains cas et empêcher l'exploitation de données, mais bien souvent, si le traitement envisagé est profitable pour la société dans son ensemble, le RGPD a prévu des solutions afin de remédier aux possibles blocages.

A. Le principe de limitation des finalités

Le RGPD impose au responsable du traitement de traiter des données à caractère personnel dans des buts spécifiques (des finalités déterminées) et de ne pas les traiter ultérieurement de manière incompatible avec ces finalités[34]. Il

29 Résumé du second alinéa de l'article 3 du RGPD proposé sur le site de l'Autorité de protection des données belge, disponible sur https://www.autoriteprotectiondonnees.be/fr/node/19237.

30 Pour plus de précisions sur le champ d'application territoriale du règlement, voy. C. DE TERWANGNE, « Définitions clés et champ d'application du RGPD », *op. cit.*, §§ 21 et s.

31 Nous avons fait le choix de nous limiter à certaines difficultés fondamentales et pas à certaines difficultés plus pratiques que les traitements de type de big data pourraient poser.

32 T. ZARSKY, « Incompatible: The GDPR in the Age of big data », *Seton Hall Law Review*, vol. 47, n° 4(2) 2017, disponible sur https://ssrn.com/abstract=3022646.

33 Groupe 29, *Déclaration du 16 septembre 2016 concernant l'impact du développement des mégadonnées sur la protection des individus à l'égard du traitement de leurs données à caractère personnel dans l'UE*, WP 221, p. 2 ; DATATILSYNET (L'autorité de protection des données norvégienne), *Big data – Privacy principles under pressure*, *op. cit.*, p. 20.

34 Art. 5, § 1er, b), du RGPD.

n'est donc pas permis de stocker des données sans raison particulière, juste « au cas où… ».

Le responsable du traitement est donc tenu de fixer ces finalités à l'avance, avant même d'entamer le traitement des données[35]. Les termes utilisés pour délimiter une finalité doivent être suffisamment précis et explicites pour circonscrire clairement le traitement de données, ce qui permettra notamment à la personne concernée d'appréhender aisément ce qui sera fait de ses données[36]. Des finalités telles que « à toutes fins utiles », « pour répondre aux nécessités de la mission », « afin notamment d'améliorer nos services »[37] restent des formulations trop vagues[38] pour répondre aux exigences de clarté et de précision exigées par le RGPD. Il faut dès lors tenter de trouver le juste milieu entre une définition trop large des finalités qui ne permet plus concrètement de cerner le sort réservé à ces données et une définition trop précise et détaillée qui s'avéra vite handicapante puisqu'elle restreindra les possibilités d'utilisation des données. Le meilleur compromis est alors, nous semble-t-il, de ne pas hésiter à énumérer plusieurs finalités, parfois proches les unes des autres, mais permettant ainsi de rester suffisamment concret, tout en offrant plus de marge de manœuvre pour le responsable du traitement[39].

Ce principe de limitation des finalités paraît facile à respecter puisqu'il suffit de savoir, avant de récolter les données, pourquoi on est amené à traiter ces données. Prenons l'exemple d'une entreprise de livraison de colis qui pose des puces GPS sur sa flotte de véhicules afin de pouvoir, à partir des données de localisation récoltées grâce à ces puces, effectuer des analyses big data dans le but d'anticiper les temps de livraison de ses colis. Dans ce cas de figure, l'entreprise n'a aucune difficulté pour définir à l'avance la finalité de son traitement, qui sera ici « d'améliorer l'anticipation des délais de livraison ».

Cependant, cela pose problème lorsqu'un responsable du traitement souhaite réutiliser des données qu'il a collectées pour une autre raison et tenter d'en tirer des informations supplémentaires, en utilisant des techniques d'analyse big data notamment[40]. Au sens du RGPD, on parlera de *traitement ultérieur des données*.

Le RGPD n'interdit pas cette réutilisation de données déjà collectées, pour autant que la finalité ultérieure soit compatible avec la finalité initiale du traitement[41]. Sont en principe

35 Groupe 29, *Opinion 03/2013 on 2 avril 2013 on purpose limitation*, WP 203, p. 15.

36 *Ibid.*, pp. 15-16.

37 Ce type de formulation était notamment utilisé par Google. À l'époque, le Groupe 29 avait exigé de la firme américaine qu'elle clarifie ce point afin de mieux informer les utilisateurs des services de Google sur l'utilisation faite des données de ceux-ci et la manière dont ces données sont partagées entre les différents services du géant de la Silicon Valley, voy. la lettre envoyée par le Groupe 29 le 23 septembre 2014 à Google, disponible sur http://ec.europa.eu/justice/article-29/documentation/other-document/files/2014/20140923_letter_on_google_privacy_policy_appendix.pdf.

38 B. Van Alsenoy *et al.*, « From social media service to advertising network: a critical analysis of Facebook's Revised Policies and Terms », 2015, disponible sur https://www.law.kuleuven.be/citip/en/news/item/facebooks-revised-policies-and-terms-v1-2.pdf ; H. Ursic et B. Custers, « Legal Barriers and Enablers to big data Reuse », *EDPL*, 2016/2, p. 213.

39 Dans ce sens, voy. la recommandation n° 20 de la Commission de la protection de la vie privée belge (l'ancien nom de l'Autorité de protection des données belge) dans son rapport sur le big data en 2016, disponible sur son internet.

40 Tel sera le cas, en gardant notre exemple d'entreprise de livraison de colis, si les puces GPS étaient déjà installées dans les véhicules pour savoir en temps réel où est situé le véhicule et éviter le vol de celui-ci. La finalité initiale du traitement (« localisation en temps réel des véhicules pour éviter les vols ») diffèrera du nouveau traitement qui ne vise pas le même objectif, même s'il réutilise les mêmes données.

41 Art. 5, § 1er, b), du RGPD.

compatibles les finalités en lien direct avec la finalité de base et que la personne concernée peut raisonnablement anticiper[42]. De plus, certaines finalités sont considérées comme étant, par nature, compatibles. Tel est notamment le cas pour les traitements ultérieurs à des fins de recherche scientifique ou à des fins statistiques[43]. Pour pouvoir bénéficier de cette compatibilité automatique, il faut alors respecter le régime spécifique prévu pour chacun de ces traitements ultérieurs[44].

Compte tenu de cette possibilité de réutiliser des données à caractère personnel pour une nouvelle finalité compatible avec la finalité initiale de la collecte des données, une entreprise qui souhaite valoriser les données qu'elle a déjà pourra le faire si les nouvelles finalités de traitement qu'elle envisage sont en lien avec la finalité pour laquelle ces données ont été collectées et que certaines mesures protectrices pour les personnes concernées sont prises par le responsable du traitement[45].

Précisons que, suite aux nouvelles opportunités offertes par les technologies big data, il faut être particulièrement vigilant au moment d'évaluer la compatibilité des deux traitements. Les technologies big data permettent en effet d'effectuer une série de choses qui ne sont pas en lien avec les finalités initiales du traitement ou qui pourraient avoir des conséquences importantes sur la personne, alors que le traitement initial ne présentait pas de risque particulier[46].

Si la réutilisation des données qui est envisagée s'éloigne trop de la raison pour laquelle celles-ci ont été récoltées, il demeure notamment possible de retourner vers la personne concernée afin d'obtenir son consentement pour de nouveaux traitements[47]. Pour être valable, ce consentement doit, pour rappel, remplir les conditions de l'article 4, 11°, du RGPD, c'est-à-dire être « une manifestation de volonté, libre, spécifique, éclairée et univoque par laquelle la personne concernée accepte, par une déclaration ou par un acte positif clair, que des données à caractère personnel la concernant fassent l'objet d'un traitement ».

Précisons qu'au moment de se lancer dans un traitement ultérieur, le responsable de traitements est tenu d'informer les personnes concernées, dans un délai raisonnable, de l'existence de ce nouveau traitement, et leur donner les informations pertinentes sur cette nouvelle finalité, conformément aux articles 13 et 14 du RGPD.

Ces limitations à la réutilisation des données ont fait dire à certains que « [l']interdiction de traitement en cas d'incompatibilité des finalités s'oppose à l'évolution d'un traitement de données qui est en quelque sorte "figé" par sa finalité réelle de départ. Si des données ont été traitées pour les besoins d'exécution d'un contrat, elles ne peuvent être traitées pour une communication à un tiers en vue d'alimenter

42 Le RGPD reprend, à son article 6, § 4, une série de critères à prendre en compte pour apprécier cette compatibilité.

43 Cons. 159 à 163 et art. 89 du RGPD et titre 4 de la loi belge du 30 juillet 2018 relative à la protection des personnes physiques à l'égard des traitements de données à caractère personnel.

44 Voy. le régime prévu au titre 4 de la loi belge du 30 juillet 2018 relative à la protection des personnes physiques à l'égard des traitements de données à caractère personnel. Le régime reste relativement souple et les mesures protectrices mentionnées sont relativement peu contraignantes.

45 Art. 6, § 4, e), du RGPD.

46 Tel est par exemple le cas pour les analyses big data permettant de cibler très précisément un individu à partir de données de contact (nom, adresse, email…), *a priori* banales, et d'ensuite lui proposer des services ultra-personnalisés en fonction de son lieu de vie. Dans ce genre de cas, le risque de discrimination est parfois trop grand pour pouvoir considérer que la réutilisation à des fins de ciblage est compatible avec la finalité initiale de la récolte de données.

47 Art. 6, § 4, du RGPD.

un processus de profilage big data, sauf consentement de la personne ou autorisation légale »[48] [49].

Certes, il est vrai que, dans certains cas, on peut considérer que ces restrictions quant à la réutilisation des données s'apparentent à une forme d'occasions manquées, mais le RGPD prévoit la plupart du temps des mécanismes permettant de dépasser ces restrictions (anonymisation des données, traitement ultérieur à des fins de recherche ou de statistique, retour vers la personne pour obtenir son consentement…), restrictions conçues pour éviter que certains soient tentés de faire n'importe quoi avec les données auxquelles ils ont accès.

B. Le principe de minimisation des données

La logique de base du big data se résume ainsi : plus on a de données (de sources diverses et variées si possible), plus les informations tirées de l'analyse de celles-ci seront précises et pertinentes.

En effet, la différence avec les autres techniques d'analyse de données réside dans le fait que les données prises individuellement n'ont que peu de valeur, mais qu'une fois regroupées pour former une quantité suffisamment conséquente de données, il sera possible de faire tourner des algorithmes sur ces données, au point d'en extraire des informations, qui auront tendance à être de plus en plus exactes au fur et à mesure que l'algorithme est alimenté en données, ce qui lui permet de s'améliorer à chaque fois qu'il reçoit un nouveau jeu de données.

Dès lors, en partant de cette logique de base, il est normal qu'un *data scientist* (spécialiste de l'analyse de données) souhaite récolter le maximum de données et des données les plus précises possible. Conscientes des capacités actuelles d'analyse de données, il est également tout à fait naturel que des organisations aient tendance à vouloir conserver les données pour pouvoir éventuellement les réutiliser plus tard[50].

Cependant, cette logique s'oppose frontalement à un des principes fondamentaux en matière de protection des données : le principe de minimisation des données[51].

En vertu de ce principe, le responsable du traitement est tenu de ne pas traiter plus de données que ce qui est nécessaire au regard des finalités pour lesquelles elles sont traitées[52]. Celui-ci doit en effet se limiter à traiter des données adéquates, pertinentes et strictement nécessaires.

Ce principe vaut tant pour la qualité des données que leur quantité. Il est en effet interdit de traiter un nombre excessif de données alors

48 Th. Léonard et D. Chaumont, « Article 6 : Licéité du traitement », GDPR-expert, disponible sur https://www.gdpr-expert.eu/article.html?id=6#difficultes probables.

49 Il avait été envisagé d'élargir cette possibilité de traiter des données à des fins incompatibles dans les cas où les intérêts légitimes du responsable ou d'un tiers primaient sur les intérêts des individus concernés. Cet élargissement aurait certes rendu moins compliqué de réutiliser des données pour « faire du big data » (voy. C. Burton *et al.*, « The Final European Union General Data Protection Regulation », *Privacy and Security Law Report*, 15 PVLR 153, 25 janvier 2016, p. 6), mais cela aurait également constitué une sérieuse exception au principe de limitation des finalités et aurait pour ainsi dire vidé celui-ci de sa substance.

50 Ce qui est potentiellement contraire au principe de limitation des finalités, si ces finalités ultérieures n'ont pas été prévues initialement (voy. *supra*, II, A).

51 Dans le cadre de cette contribution, nous ne ferons pas de distinction entre le principe de minimisation (relatif à la nature des données et au volume des données, art. 5, § 1er, c), du RGPD) et de limite de conservation (relatif à la durée de conservation des données, art. 5, § 1er, e), du RGPD). Ces deux principes découlent en effet de la même idée qu'il faut éviter de traiter des données à caractère personnel quand cela est (devenu) inutile.

52 Art. 5, § 1er, c), du RGPD.

que cela s'avère parfaitement disproportionné. Par ailleurs, certaines données particulièrement sensibles (données de santé…), même en faible quantité, ne peuvent être traitées si le responsable du traitement peut parvenir à ses objectifs sans recourir à ce genre de données et peut se contenter de données moins intrusives.

Ce principe impose également au responsable de ne traiter des données à caractère personnel que dans les cas où il ne peut faire raisonnablement autrement pour atteindre les finalités qu'il poursuit[53]. Ainsi, si ce dernier a la possibilité d'y parvenir avec des données anonymisées, ou pseudonymisées[54], il est tenu de se limiter à ce type de données, quitte à lui-même faire usage de techniques d'anonymisation permettant, en amont de tout traitement, d'éviter de récolter des données personnelles inutiles[55] [56].

De plus, le responsable du traitement ne peut continuer à conserver des données plus longtemps que nécessaire et devra donc les effacer (et/ou les anonymiser) une fois qu'il aura atteint la finalité pour laquelle elles ont été récoltées[57].

Notons qu'il est autorisé de continuer à stocker les données après avoir atteint la finalité initiale du traitement et ainsi conserver les données « pour des durées plus longues dans la mesure où elles seront traitées exclusivement à des fins archivistiques dans l'intérêt public, à des fins de recherche scientifique ou historique ou à des fins statistiques conformément à l'article 89, paragraphe 1, pour autant que soient mises en œuvre les mesures techniques et organisationnelles appropriées requises par le présent règlement afin de garantir les droits et libertés de la personne concernée »[58].

Compte tenu du principe de minimisation des données, il est nécessaire pour les *data scientists* d'apprécier, pour chaque nouveau jeu de données qu'ils envisagent de récolter ou de réutiliser, si ces données sont nécessaires, simplement utiles, ou d'un intérêt minime (dans quel cas ils ne pourront pas les (ré)utiliser)[59]. Après cette première question, ils devront également s'assurer qu'ils ne peuvent pas effectuer ses analyses avec des données « moins intrusives » (données anonymes, données pseudonymisées, données non sensibles…).

Cette prise en compte du principe de minimisation de données peut sembler faire obstacle au développement du big data en Europe et pourrait freiner certaines recherches. Cependant, c'est oublier que ce principe de minimisation des données n'empêche pas de traiter des données à caractère personnel si celles-ci sont nécessaires. Il force uniquement chaque responsable du traitement à se poser la question de la nécessité des données récoltées et à justifier sa réponse, afin de se conformer au principe d'« accountability » qui impose à tout responsable de traitements de pouvoir prouver

53 Cons. 39 du RGPD.

54 Données initialement directement identifiantes (un nom) ayant été remplacées par un code ne permettant plus cette réidentification directe. Il reste cependant possible, si nécessaire, de réidentifier la personne grâce à la table de concordance (nom/numéro).

55 Mesures imposées quand cela est possible, notamment en vertu du principe de « privacy by design » qui oblige les responsables de traitement à tenir compte des principes du RGPD au moment de concevoir de futurs traitements de données, voy. art. 25 et cons. 28 et 29 du RGPD.

56 Comme rappelé précédemment, il devient de plus en plus compliqué d'anonymiser complètement des données à caractère personnel, voy. *supra*, I, A.

57 Art. 5, § 1er, e) et cons. 39 du RGPD.

58 Art. 5, § 1er, e), du RGPD. À ce sujet, voy. la section précédente au sujet des finalités ultérieures compatibles.

59 Précisons que le volume de données doit être suffisant pour que l'algorithme ne produise pas de données à caractère personnel erronées, voy. *infra,* C.

qu'il respecte bien le RGPD et en l'occurrence le principe de minimisation des données[60].

C. Le principe d'exactitude des données

Le troisième principe fondamental du RGPD qui pourrait poser des difficultés est le principe d'exactitude des données qui exige du responsable du traitement, comme son nom l'indique, de ne traiter que des données à caractère personnel, exactes et mises à jour, si cela s'avère nécessaire[61].

Cette obligation reste cependant une obligation de moyen, puisque le RGPD précise que le responsable du traitement se doit de prendre « toutes les mesures raisonnables […] pour que les données à caractère personnel qui sont inexactes, eu égard aux finalités pour lesquelles elles sont traitées, soient effacées ou rectifiées sans tarder »[62].

Seules les données *objectivement* inexactes doivent être corrigées ou supprimées. Ainsi, un jugement de valeur, un avis subjectif ne peut être considéré comme inexact, sauf s'il est lui-même fondé sur des données objectivement inexactes[63].

La personne concernée a d'ailleurs le droit de faire rectifier ces informations erronées et « d'obtenir que les données à caractère personnel incomplètes soient complétées »[64] afin d'ajouter certains éléments qui peuvent apporter un éclairage différent aux données initialement collectées.

Dans le cadre de traitements de données de type big data, l'appréciation du caractère exact des données est particulièrement délicat et doit se faire à plusieurs niveaux[65] afin d'éviter de produire des données à caractère personnel erronées, ce qui peut s'avérer particulièrement problématique lorsqu'une décision automatisée est prise en se fondant notamment sur cette erreur, ce qui peut avoir pour conséquence, par exemple, que la personne concernée peut se voir refuser un prêt hypothécaire au motif que son risque d'insolvabilité est erronément élevé.

Le premier niveau auquel il faut veiller est l'exactitude des données constituant le set de données utilisé.

Ensuite, il faut s'assurer que l'algorithme produit des résultats corrects et non biaisés. En effet, les résultats obtenus ne sont pas toujours parfaits et il peut arriver, par exemple, que certaines personnes soient classées à tort dans une catégorie à laquelle elles ne devraient pas objectivement appartenir (fraudeur potentiel) ou qu'une catégorie de personnes soit systématiquement affectée erronément à une classe déterminée (profil à haut risque). Cette mauvaise classification peut provoquer la stigmatisation de certains profils d'individu et amener à des discriminations parfois involontaires de la part du responsable du traitement.

Ce type d'erreur peut survenir lorsque le set de données ayant servi à entraîner le modèle algorithmique est inadapté, insuffisant ou lorsque le set de base est mal calibré de sorte qu'il existe des biais dans les données ayant servi à l'apprentissage[66]. Des résultats peu

60 Art. 5, § 2 et 24 du RGPD. À ce sujet, nous renvoyons à notre contribution « Les obligations générales du responsable du traitement et la place du sous-traitant », in *Le Règlement général sur la protection des données : Analyse approfondie*, *op. cit.*, à paraître.

61 Art. 5, § 1er, d), du RGPD.

62 *Idem*.

63 C. de Terwangne, « Les principes relatifs au traitement des données à caractère personnel et à sa licéité », in *Le Règlement général sur la protection des données : Analyse approfondie*, *op. cit.*, § 27.

64 Art. 16 du RGPD.

65 Pour une étude plus complète de la question, voy. entre autres le rapport sur le big data de la Commission de la protection de la vie privée réalisé en 2016, précité, pp. 22 et s.

66 Ces biais peuvent parfois résulter d'une mauvaise méthodologie dans le choix des données ou la manière d'entraîner le modèle algorithmique, voire

probants peuvent également découler du fait que l'algorithme ait été mal paramétré, voire mal choisi. Il est donc important de vérifier que l'algorithme produit des résultats suffisamment probants et non biaisés avant de l'utiliser.

Enfin, les résultats d'une analyse big data doivent être interprétés correctement. Les techniques big data permettent de trouver des corrélations entre différents éléments. Cette corrélation pourra notamment servir ensuite pour classer certains individus dans des catégories particulières. Une corrélation entre ces deux éléments ne doit cependant pas être confondue avec un lien de cause à effet entre ceux-ci. Cela signifie uniquement que statistiquement, les deux éléments ont un lien (parfois indirect). Néanmoins, cela ne permet pas d'en déduire une réalité objective[67]. Cela ne peut servir que d'indice, à défaut de quoi, ce type de procédé provoquera une forme de déterminisme dicté par le profil numérique de chaque individu[68].

Le principe d'exactitude des données empêche donc que des décisions pouvant éventuellement avoir des effets importants dans la vie des gens soient prises, parfois automatiquement, sur base de données erronées ou via un processus non fiable ou biaisé.

Cependant, le fonctionnement de certains types d'algorithmes (les réseaux neuronaux notamment) étant difficilement compréhensible par un humain[69], il reste dans certains cas compliqué de s'assurer de la fiabilité des résultats.

D. Les droits des personnes concernées

Le RGPD accorde de nombreux droits aux personnes concernées afin que celles-ci sachent ce qui est fait de leurs données et puissent dans une certaine mesure contrôler leur utilisation par le responsable de traitements.

Dans un contexte big data, deux droits en particulier méritent d'être évoqués : le droit à l'information et le droit de ne pas faire l'objet d'une décision automatisée[70].

Les deux droits que nous avons retenus s'appliquent spécifiquement aux traitements de données consistant en *une prise de décision automatisée*[71], c'est-à-dire une décision prise « par des moyens technologiques sans intervention humaine »[72].

1. Le droit à l'information

La personne concernée doit recevoir de la part du responsable du traitement une série d'informations sur le traitement qui sont reprises à

de l'existence de préjugés cachés dans les données (surreprésentation d'une catégorie de la population dans le set de données de base, par exemple, de sorte que l'algorithme aura tendance à discriminer les personnes membres de cette catégorie ou celles n'appartenant pas à celle-ci). Ce phénomène est appelé « distorsion des données ». Sur le sujet, voy. notamment K. Crawford, « The hidden biases in big data », *Harvard Business review*, 2013, disponible sur https://hbr.org/2013/04/the-hidden-biases-in-big-data.

67 Ce n'est pas, par exemple, parce qu'un algorithme a établi qu'il y a statistiquement de fortes chances qu'une personne fraude qu'il faut automatiquement la classer comme fraudeuse. Un examen détaillé et individuel est nécessaire pour en arriver à cette conclusion.

68 Sur le sujet, voy. en particulier les travaux de A. Rouvroy et récemment « Homo juridicus est-il soluble dans les données ? », in *Droit, normes et libertés dans le cybermonde*, coll. du CRIDS, Bruxelles, Larcier, 2018, pp. 417 et s.

69 Voy. *infra*, D, 1.

70 D'autres droits auraient pu également être étudiés. Nous pensons notamment au droit à la portabilité des données prévu à l'article 20 du RGPD.

71 Ces types de traitement nous intéressent particulièrement puisqu'ils sont généralement fondés sur des analyses big data de profilage.

72 Groupe 29, *Lignes directrices relatives à la prise de décision individuelle automatisée et au profilage aux fins du règlement (UE) 2016/679*, version révisée et adoptée le 6 février 2018, WP 251 rev.01.

l'article 13 du RGPD si la récolte des données est faite directement auprès de la personne concernée et à l'article 14 si elles ont été collectées indirectement[73].

Parmi les informations à communiquer à la personne concernée, est notamment reprise « l'existence d'une prise de décision automatisée, y compris un profilage, visée à l'article 22, paragraphes 1 et 4, et, au moins en pareils cas, des informations utiles concernant la logique sous-jacente, ainsi que l'importance et les conséquences prévues de ce traitement pour la personne concernée »[74].

La personne aura donc le droit de comprendre la logique sous-jacente à la prise de décision. Cela ne signifie donc pas qu'elle doit forcément avoir accès à l'algorithme utilisé, mais bien qu'on lui explique quelles données sont utilisées et comment l'algorithme fonctionne, en des termes simples et compréhensibles[75].

Cette démarche d'explicabilité permet de pallier le manque de transparence qui peut exister pour les traitements utilisant des algorithmes d'une grande complexité.

Cependant pour certaines technologies d'intelligence artificielle, le fonctionnement de l'algorithme demeure difficilement intelligible même pour les experts, c'est ce qu'on appelle le « black box paradox » : on connaît les données qui sont soumises à l'algorithme et on découvre ce qui en ressort, sans véritablement comprendre ce qui se passe à l'intérieur de la « boîte ».

Dès lors, il devient particulièrement compliqué pour un responsable du traitement d'expliquer comment son algorithme fonctionne et pourquoi il en arrive à ce résultat. Le problème est ici donc avant tout technique. Le droit à une explication devient ainsi un véritable défi pour les spécialistes du big data[76].

2. *Le droit de ne pas faire l'objet d'une décision automatisée*

Pour éviter qu'un être humain ne se sente soumis exclusivement à une machine, ce qui peut s'avérer particulièrement dégradant, le RGPD prévoit que « la personne concernée a le droit de ne pas faire l'objet d'une décision fondée exclusivement sur un traitement automatisé, y compris le profilage, produisant des effets juridiques la concernant ou l'affectant de manière significative de façon similaire »[77]. Le considérant 71 du RGPD cite comme exemple de pareille décision « le rejet automatique d'une demande de crédit en ligne ou des pratiques de recrutement en ligne sans aucune intervention humaine ».

Ce droit ne s'applique toutefois pas, si ce type de prise de décision est autorisé légalement, si la personne concernée a explicitement consenti à cette méthode, ou lorsque ceci est nécessaire à l'exécution d'un contrat entre le responsable du traitement et la personne concernée[78] [79].

73 Ce droit découle notamment du principe général de transparence contenu à l'art. 5, § 1er, a), du RGPD.

74 Art. 13, § 2, f) et 14, § 2, g), du RGPD.

75 Art. 12 et cons. 58 du RGPD.

76 Sur le problème du « black box paradox » dans le contexte du RGPD, voy. notamment S. Wachter, B. Mittelstadt et Chr. Russell, « Counterfactual Explanations without Opening the Black Box: Automated Decisions and the GDPR », *Harvard Journal of Law & technology*, 2018, pp. 841 et s.; M. Brkan, « Do Algorithms Rule the World? Algorithmic Decision-Making in the Framework of the GDPR and Beyond », disponible sur https://ssrn.com/abstract=3124901.

77 Art. 22, § 1er, du RGPD.

78 Art. 22, § 2, du RGPD.

79 Dans ces trois cas, des mesures appropriées doivent être prévues afin de pallier les risques de pareils processus. Dans les deux dernières hypothèses, la personne a même le droit « d'obtenir une intervention humaine de la part du responsable du traitement, d'exprimer son point de vue et de contester la décision », voy. art. 22, § 3, du RGPD.

Les analyses big data servant bien souvent de fondement à une décision automatisée, dès lors ce droit offre la possibilité d'éviter d'être catégorisé (profilé) par un algorithme et ensuite soumis à une décision automatisée, ou en tout cas de faire valoir son point de vue auprès d'une personne humaine.

CONCLUSION

Depuis maintenant quelques années déjà de nombreuses entreprises et États ont commencé à développer des stratégies big data censées permettre de tirer un maximum d'informations des données à disposition, sans parfois tenir compte de la législation relative à la protection des données qui existait à l'époque.

Avec l'entrée en application du RGPD, une partie de ceux-ci ont découvert l'existence de certaines règles en matière de protection des données qui existaient déjà depuis près de trente ans et que le RGPD n'a pas fondamentalement modifiées. Cependant, contrairement à ce qui pouvait exister sous l'empire de la directive, le non-respect du RGPD n'est plus une option viable au vu des montants pharaoniques prévus pour les amendes administratives.

Ils ont donc été, pour certains, étonnés de découvrir que la plupart du temps les traitements de type big data envisagés sont amenés à contenir des données à caractère personnel. La définition de la notion étant à ce point large que, bien souvent, derrière des données *a priori* anodines (de simples chiffres) se cachaient des données à caractère personnel.

Dans une perspective de big data, la présence de données à caractère personnel insoupçonnée est d'autant plus probable que les capacités d'analyse, de croisement de données sont tellement évoluées aujourd'hui et les sources de données disponibles tellement plus nombreuses qu'auparavant qu'une donnée autrefois anonyme, ou anonymisée, peut (re)devenir une donnée à caractère personnel, ce qui aura pour conséquence que les règles du RGPD viennent potentiellement restreindre les possibilités de réutilisation des données déjà récoltées.

Le RGPD prévoit en effet une série de principes qui semblent opposés à la logique même du big data.

Le big data vise notamment à profiter des données déjà collectées pour les réutiliser à d'autres fins (pour en déduire des tendances par exemple ou pour essayer de profiler au mieux ses clients). Le RGPD quant à lui prévoit que les données à caractère personnel ne peuvent être traitées que pour des finalités définies à l'avance et qu'on ne peut s'en éloigner trop fortement.

Le big data est fondé sur le principe selon lequel plus il y a de données, meilleurs seront les résultats, là où le RGPD impose de récolter le minimum de données possible.

Le big data ne permet pas d'obtenir des données exactes à tous les coups (voire ne permet même pas de le vérifier), alors que le RGPD interdit de traiter (collecter, stocker ou produire) des données erronées.

Il est de plus parfois difficile d'expliquer aux personnes concernées, dans des termes simples, la logique sous-jacente à un traitement de type big data.

Le big data et le RGPD pourraient donc être vus comme deux termes inconciliables, ce qui aurait pour conséquence de voir l'Europe rater le tournant actuel de l'économie du numérique.

Ce n'est pourtant pas le cas, le RGPD n'empêche pas de « faire du big data », il ne fixe que certains principes qui ont vocation à protéger les intérêts de chacun, à défaut de quoi une méfiance dans le chef des citoyens, des

personnes concernées pour reprendre le vocabulaire du RGPD, risque de s'installer et alors tout le monde hésitera à confier ses données à une entreprise, ce qui n'est pas forcément souhaitable pour le bien de la société qui a besoin d'un climat de confiance[80].

Pour s'en convaincre, il suffit de voir les débats actuels aux États-Unis où de nombreuses personnes quittent Facebook au motif qu'ils n'ont plus confiance en cette entreprise[81] et où certains plaident pour un cadre légal plus protecteur.

Dans cette optique, au lieu de percevoir le RGPD comme un instrument juridique inutilement restrictif, ses détracteurs peuvent également le concevoir comme un mal nécessaire, ou un investissement dans la relation de confiance entre responsable du traitement et personne concernée, pour parler en des termes plus positifs.

80 V. Jourová (EU Commissioner for Justice, Consumers and Gender Equality), « The EU Data Protection Reform and big data », Factsheet January 2016, disponible sur http://ec.europa.eu/newsroom/just/document.cfm?doc_id=41523.

81 Voy. A. Perrin, « Americans are changing their relationship with Facebook », disponible sur http://www.pewresearch.org/fact-tank/2018/09/05/americans-are-changing-their-relationship-with-facebook/.