EUROPEAN DATA PROTECTION SUPERVISOR

Avis n° 7/2015

Relever les défis des données massives

Un appel à la transparence, au contrôle par l'utilisateur, à la protection des données dès la conception et à la reddition de comptes

19 novembre 2015

Le contrôleur européen de la protection des données (CEPD) est une institution indépendante de l'UE. Le contrôleur est chargé en vertu de l'article 41, paragraphe 2, du règlement (CE) n° 45/2001, «[e]n ce qui concerne le traitement de données à caractère personnel… de veiller à ce que les libertés et droits fondamentaux des personnes physiques, notamment leur vie privée, soient respectés par les institutions et organes communautaires», et «… de conseiller les institutions et organes communautaires et les personnes concernées pour toutes les questions concernant le traitement des données à caractère personnel».

Le contrôleur européen et le contrôleur adjoint ont été nommés en décembre 2014 avec comme mission spécifique d'être plus constructifs et proactifs et ils ont publié en mars 2015 une stratégie quinquennale exposant la manière dont ils entendaient mettre en œuvre ce mandat et en rendre compte.

Cet avis fait suite à l'avis précédent du CEPD sur l'éthique numérique[1]. Dans le présent avis, le CEPD aborde le défi du passage en «mode numérique» de la protection des données – le troisième objectif de la stratégie du CEPD – en «visant à adapter les principes de protection des données au monde numérique», compte tenu également des projets de l'UE concernant le marché unique numérique. L'avis est conforme à l'approche du groupe de travail «Article 29» sur les aspects liés à la protection des données de l'utilisation des nouvelles technologies, comme l'«Internet des objets», à laquelle le CEPD a contribué en tant que membre à part entière du groupe de travail.

BIG DATA	DONNÉES MASSIVES
TRANSPARENCY	TRANSPARENCE
USER CONTROL	CONTRÔLE DES UTILISATEURS
DIGNITY	DIGNITÉ
PRIVACY BY DESIGN	RESPECT DE LA VIE PRIVÉE DÈS LA CONCEPTION
ACCOUNTABILITY	REDDITION DE COMPTES
FREEDOM	LIBERTÉ
PRIVACY	RESPECT DE LA VIE PRIVÉE

«Le droit d'être laissé tranquille est effectivement le début de toute liberté»[2].

Les données massives, si elles sont traitées de façon responsable, peuvent apporter des avantages et des gains d'efficacité significatifs à la société et aux personnes physiques, non seulement dans les domaines de la santé, de la recherche scientifique et de l'environnement, mais aussi dans d'autres domaines spécifiques. Cependant, les effets réels et potentiels du traitement d'énormes quantités de données sur les droits et les libertés des personnes concernées, y compris leur droit au respect de la vie privée, suscitent de graves inquiétudes. **Les défis et les risques liés aux données massives requièrent donc une protection plus efficace des données.**

La technologie ne devrait pas nous dicter nos valeurs et nos droits, mais la promotion de l'innovation et la sauvegarde des droits fondamentaux ne doivent pas pour autant être perçues comme incompatibles avec la technologie. De nouveaux modèles commerciaux exploitant de nouvelles capacités de collecte massive de données, de transmission instantanée, de combinaison et de réutilisation de données à caractère personnel pour des finalités imprévues, ont fait peser de nouvelles contraintes sur les principes relatifs à la protection des données, qui requièrent un examen approfondi de la façon dont ils sont appliqués.

La législation européenne relative à la protection des données a été élaborée pour protéger nos valeurs et nos droits fondamentaux, y compris notre droit au respect de la vie privée. **La question n'est pas de savoir *si* la législation relative à la protection des données doit être appliquée aux données massives («*big data*»), mais *comment* l'appliquer de façon innovante dans de nouveaux environnements.** Nos principes actuels en matière de protection des données, notamment la transparence, la proportionnalité et la limitation de la finalité, constituent la ligne de référence dont nous aurons besoin pour protéger nos droits fondamentaux de façon plus dynamique dans le monde des données massives. Ils doivent toutefois être complétés par de «nouveaux» principes, qui se sont développés au fil des années, comme la reddition de comptes et le respect de la vie privée dès la conception et par défaut. Le paquet sur la réforme de la protection des données dans l'UE doit renforcer et moderniser le cadre réglementaire[3].

L'UE entend optimiser la croissance et la compétitivité en exploitant les données massives. Mais le marché unique numérique ne peut pas importer, sans faire preuve d'un esprit critique, les technologies et les modèles commerciaux fondés sur les données, qui sont devenus le principal courant économique dans d'autres régions du monde. Il doit, en revanche, apparaître comme le chef de file du développement d'un traitement responsable des données à caractère personnel. L'internet a évolué de telle manière que la surveillance – le suivi du comportement des personnes – est considérée comme le modèle de revenus indispensable pour certaines entreprises parmi les plus performantes. Cette évolution appelle une évaluation critique et la recherche d'autres options.

Quoi qu'il en soit et indépendamment des modèles commerciaux retenus, les organisations qui traitent des volumes importants d'informations personnelles doivent se conformer à la législation applicable en matière de protection des données. Le contrôleur européen de la protection des données (CEPD) considère qu'un développement responsable et durable des données massives doit reposer sur **quatre éléments essentiels:**

- les organisations doivent être beaucoup plus **transparentes** en ce qui concerne la manière dont elles traitent les données à caractère personnel;

- elles doivent donner aux utilisateurs un degré accru de **contrôle** sur la manière dont leurs données sont utilisées;
- elles doivent **intégrer** une protection des données conviviale dans leurs produits et services et
- elles doivent rendre des **comptes** sur ce qu'elles font.

En ce qui concerne la **transparence**, les personnes concernées doivent recevoir des informations claires sur les données qui sont traitées, notamment les données observées ou déduites les concernant, être mieux informées sur la manière dont leurs données sont utilisées et sur les finalités pour lesquelles elles sont utilisées, y compris la logique algorithmique qui sert à déterminer les hypothèses et les prévisions à leur sujet.

Le contrôle par l'utilisateur contribuera à donner plus d'autonomie aux personnes concernées pour mieux déceler les préjugés injustes et contester les erreurs. Il permettra d'empêcher l'utilisation secondaire des données pour des finalités qui ne sont pas conformes aux attentes légitimes des personnes concernées. Grâce à un contrôle par l'utilisateur de nouvelle génération, les personnes concernées disposeront, le cas échéant, d'un choix plus authentique et mieux informé et auront davantage de possibilités de mieux utiliser leurs données personnelles.

Un **droit d'accès solide et un droit à la portabilité des données ainsi que des mécanismes efficaces de retrait** («*opt-out*») peuvent constituer une condition préalable pour permettre aux utilisateurs de mieux contrôler leurs données et également contribuer à la mise au point de nouveaux modèles commerciaux et à une utilisation plus transparente et efficace des données à caractère personnel.

En intégrant la protection des données dans la conception de leurs systèmes et processus et **en ajustant la protection des données pour permettre une transparence et un contrôle plus véritables par les utilisateurs, les responsables du traitement tenus de rendre des comptes** pourront également bénéficier des avantages des données massives, tout en veillant au respect de la dignité et des libertés des personnes.

La protection des données ne constitue toutefois qu'une partie de la réponse. L'UE doit déployer de façon plus cohérente les outils modernes disponibles, notamment dans le domaine **de la protection des consommateurs, de la législation antitrust, de la recherche et du développement**, pour garantir la protection et le choix sur le marché, où des services respectueux de la vie privée pourront prospérer.

Pour relever les défis que posent les données massives, nous devons **permettre l'innovation tout en protégeant les droits fondamentaux**. Il appartient désormais aux entreprises et aux autres organisations qui déploient d'importants efforts dans la recherche de solutions innovantes pour l'utilisation des données à caractère personnel de faire preuve du même esprit innovant dans la mise en œuvre de la législation relative à la protection des données.

Prenant appui sur les contributions antérieures d'universitaires et de nombreux régulateurs et parties prenantes, **le CEPD veut favoriser une discussion nouvelle, ouverte et éclairée tant à l'intérieur qu'à l'extérieur de l'UE**, en faisant participer davantage la société civile, les concepteurs, les entreprises, le monde académique, les pouvoirs publics et les régulateurs à l'élaboration de la meilleure façon d'exploiter le potentiel créatif de l'industrie pour faire appliquer la loi et protéger notre vie privée et nos autres droits fondamentaux le mieux possible.

TABLE DES MATIÈRES

1. Analyse des données massives: opportunités, risques et défis

1.1 «Données massives» et «analyse des données massives»

De façon générale, le dénominateur commun de toutes les définitions disponibles veut que les «données massives»[4] fassent référence à la pratique qui consiste à combiner d'énormes volumes d'informations provenant de différentes sources et à les analyser en utilisant des algorithmes plus complexes pour mieux étayer les décisions. Les données massives dépendent non seulement de la capacité accrue de la technologie à supporter la collecte et le stockage de grandes quantités de données, mais également de sa capacité à analyser, comprendre et exploiter pleinement la valeur de ces données (en particulier, grâce à des applications d'analyse).

Les données massives font espérer qu'elles puissent conduire, en fin de compte, à des décisions meilleures et plus éclairées. Les données massives peuvent, par exemple, conduire à une meilleure vision de la recherche scientifique et médicale, à une meilleure connaissance personnelle des particuliers, des produits, des services et des traitements médicaux, qui sont plus personnalisés et mieux adaptés à la personne, et à de meilleures décisions automatisées pour les entreprises et les 'autres organisations qui traitent les données. Ces décisions automatisées peuvent, à leur tour, aboutir à une efficacité accrue, avec des promesses de diverses applications commerciales et autres.

Les informations traitées par les applications reposant sur des données massives n'ont pas toujours un caractère personnel: les données générées par des capteurs à des fins de surveillance des phénomènes naturels ou atmosphériques, comme le temps ou la pollution, ou à des fins de surveillance des aspects techniques de procédés de fabrication, peuvent ne pas se rapporter à «*une personne physique identifiée ou identifiable*». Mais l'une des valeurs les plus importantes des données massives pour les entreprises et les gouvernements découle de la surveillance des comportements *humains*, aux niveaux collectif et individuel, et réside dans leur potentiel prédictif.

L'une des conséquences est l'émergence d'un modèle de revenus pour les sociétés de l'Internet, qui repose sur le suivi de l'activité en ligne. Ces «données massives» devraient être considérées comme des données à caractère personnel, même dans les cas où des techniques d'anonymisation ont été utilisées: il est de plus en plus facile de déduire l'identité d'une personne de la combinaison de données prétendument «anonymes» et d'autres informations accessibles au public, par exemple sur les médias sociaux. Par ailleurs, avec l'avènement de l'«Internet des objets», une grande partie des données collectées et communiquées par le nombre croissant de dispositifs et de capteurs personnels et autres seront des données à caractère personnel: les données qu'ils recueillent peuvent aisément être mises en rapport avec les utilisateurs de ces dispositifs, dont le comportement sera surveillé. Elles peuvent englober des données extrêmement sensibles, notamment des informations sur la santé et des renseignements concernant nos modes de pensée et notre psychologie.

Les applications reposant sur des données massives qui traitent des données à caractère personnel évaluent souvent certains aspects des personnes, comme la santé ou les risques financiers. Dans d'autres cas, les entreprises utilisent des données massives pour commercialiser de façon plus efficace et efficiente des produits ou des services qui nous sont destinés et/ou fournir un service plus personnalisé. Un nombre croissant d'autres applications évaluent également les personnes physiques à diverses fins: un employeur peut tirer parti de données massives pour présélectionner les candidats les plus prometteurs pour un poste et les

voyageurs peuvent utiliser des applications pour voir quelle compagnie de taxis ou quelle chambre d'hôte offre le meilleur service. Dans d'autres cas encore, les organisations ont besoin de nos données pour mener diverses études et souhaitent dégager des tendances générales et des corrélations dans les données[5].

1.2 Quels sont aujourd'hui les principaux risques et défis que posent les données massives?

L'application des données massives présente des avantages considérables pour les particuliers et la société, mais elle soulève aussi de graves inquiétudes concernant son impact potentiel sur la dignité et les droits et libertés des personnes concernées, notamment leur droit au respect de la vie privée. Ces risques et défis ont déjà fait l'objet d'analyses multiples par des spécialistes de la protection des données du monde entier[6] et le CEPD se bornera donc à ne relever que quelques préoccupations majeures.

Manque de transparence. Tandis que la complexité du traitement des données augmente, les organisations réclament souvent le secret concernant la manière dont les données sont traitées pour des raisons de confidentialité commerciale. Comme l'indiquait un rapport de la Maison Blanche de 2014, «*certains des défis les plus importants révélés par cet examen concernent la façon dont une analyse des données massives peut... créer un environnement décisionnel si opaque que l'autonomie individuelle disparaît dans une série impénétrable d'algorithmes*»[7]. À moins que les personnes physiques ne reçoivent les informations appropriées et ne disposent d'un contrôle adéquat, elles «*seront soumises à des décisions qu'elles ne comprennent pas et sur lesquelles elles n'ont aucun contrôle*»[8]. Les particuliers ne peuvent pas exercer un contrôle efficace sur leurs données et donner un consentement éclairé lorsque celui-ci est requis. C'est d'autant plus vrai que les finalités futures précises de toute utilisation secondaire des données peuvent ne pas être connues au moment de la collecte des données. Dans ce cas, les responsables du traitement peuvent ne pas être en mesure d'expliquer aux personnes concernées ce qu'il va advenir précisément de leurs données et d'obtenir leur consentement, si nécessaire, ou être réticents à le faire.

Le **déséquilibre d'information** entre les organisations qui détiennent les données et les personnes concernées dont elles traitent les données risque de s'accroître avec le développement des applications reposant sur des données massives[9].

Ne pas résoudre ces problèmes peut faire courir le risque d'une atteinte aux principes fondamentaux de la protection des données. Les opportunités que laissent entrevoir les données massives incitent à collecter autant de données que possible et à les conserver aussi longtemps que possible pour des finalités futures encore inconnues. Certains partisans des données massives réclament des dérogations aux principes de base, en particulier ceux de la limitation de la finalité et de la minimisation des données, et allèguent que ces principes ne devraient pas (ou pas totalement) s'appliquer au traitement des données massives. Les données massives remettent également en question les principes d'exactitude et de pertinence des données. Les applications reposant sur des données massives visent généralement à collecter des données auprès de diverses sources, sans procéder à une vérification approfondie de la pertinence ou de l'exactitude des données ainsi collectées.

L'une des utilisations potentiellement les plus importantes des données massives est de prédire ce qui va probablement se produire, mais ne s'est pas encore produit, et ce que nous allons probablement faire, mais n'avons pas encore fait. Ainsi, les données massives pourraient être utilisées pour prédire les résultats d'un enfant à l'école ou la vulnérabilité d'un

adulte aux maladies ou à un décès prématuré, à un défaut de remboursement de crédit ou à la perpétration d'un crime. En dépit des bénéfices potentiels, un commentateur a décrit les données comme le «problème de pollution de l'ère de l'information»[10], avec un risque de «**dictature des données**», dans laquelle, selon une étude d'une autorité européenne de protection des données, «*nous ne sommes plus jugés sur nos actes, mais sur la base de ce que toutes les données nous concernant indiquent que pourraient être nos actes*»[11].

Les avantages escomptés d'une prédiction fondée sur des statistiques peuvent renforcer davantage une confiance excessive dans ses capacités. **Les applications reposant sur des données massives peuvent trouver des corrélations trompeuses dans les données**, même lorsqu'il n'existe pas de lien direct de causalité entre deux phénomènes qui révèlent une corrélation étroite. Dans de tels cas, il existe un risque de tirer des conclusions erronées, mais également – lorsqu'elles sont appliquées au niveau individuel – **potentiellement injustes et discriminatoires**.

Ces caractéristiques et d'autres des données massives, l'utilisation étendue de décisions automatisées et l'analyse prédictive peuvent également conduire à des changements plus larges et indésirables dans l'évolution de nos sociétés. Surtout, elles peuvent mener à une discrimination, à un renforcement des stéréotypes existants, ainsi qu'à une **ségrégation et à une exclusion culturelles et sociales**[12].

Il est possible d'accumuler d'énormes ensembles de données à caractère personnel pour alimenter l'analyse des données massives en raison du suivi constant et invisible de l'activité en ligne. Cette surveillance peut aussi avoir un **effet dissuasif sur la créativité et l'innovation**.

Les analyses de données massives servent à identifier un comportement qui, statistiquement parlant, présente moins de risque et produit plus de valeur pour les organisations qui traitent les données. On observe une tendance à décourager ou à pénaliser la spontanéité, l'expérimentation ou la déviation par rapport à la «norme» statistique et à récompenser un comportement conformiste. Ainsi, les secteurs de la banque et de l'assurance ont un intérêt évident à avoir une vision approfondie du risque posé par un particulier qui pourrait être révélé en combinant des ensembles de données générés par son activité sur les médias sociaux et par des dispositifs connectés suivant la localisation et d'autres données personnelles et le nombre croissant d'objets connectés. La nécessité d'obtenir un prêt ou une couverture d'assurance pourrait pousser ou contraindre des individus à éviter le contact avec certaines personnes ou entreprises ou à visiter des quartiers où les taux de criminalité sont élevés de la même manière que des personnes sont incitées à installer des «boîtes noires» qui permettent à un responsable de traitement externe de les contrôler pendant qu'elles conduisent[13].

Le fait même que notre comportement soit constamment suivi et analysé peut également nous inciter à faire attention à la manière dont nous nous comportons et nous encourage à nous conformer, à l'avance, à ce que nous percevons comme étant la norme escomptée. Ces tendances peuvent également avoir un effet paralysant sur la liberté d'expression et sur d'autres activités nécessaires à la préservation d'une société démocratique, comme l'exercice du droit à la liberté de réunion et d'association.

Vus sous cet angle, les droits au respect de la vie privée et à la protection des données à caractère personnel sont une condition préalable pour que les personnes concernées puissent développer leur personnalité et vivre leur vie en tant qu'être humain indépendant, une

condition préalable à l'exercice des droits et libertés les plus chers ainsi qu'**une condition préalable pour que les personnes et la société innovent**.

Pour assurer la protection et le maintien de nos valeurs et nos droits fondamentaux, y compris notre capacité à continuer à innover en tant que société et en tant qu'individus, les données massives doivent être utilisées de façon plus responsable et durable. Comme indiqué dans notre avis n° 4/2015 précédent, il est urgent de créer un écosystème de protection des données massives composé:

- d'organisations beaucoup plus **transparentes** en ce qui concerne la manière dont elles traitent les données à caractère personnel;
- de personnes capables de tirer parti d'un degré accru de **contrôle** sur la manière dont leurs données sont utilisées;
- d'une protection des données intégrée dans la conception des produits et des services et
- de responsables du traitement davantage tenus de rendre des comptes.

Chacun de ces quatre points sera brièvement abordé dans le présent avis.

2. Transparence: mettre un terme au profilage clandestin

2.1 Faire connaître la logique qui sous-tend l'analyse des données massives

La transparence des décisions automatisées joue un rôle de plus en plus important dans l'avènement de l'analyse des données massives. Faire connaître la logique qui sous-tend le processus décisionnel peut aider les particuliers à mieux vérifier l'exactitude et l'équité des conclusions tirées par les organisations qui traitent les données et affectent les individus. Ils pourront mieux comprendre et peut-être rectifier les critères sous-jacents et les facteurs qui influencent la décision.

En tant que société, nous devons être capables de regarder dans la «boîte noire» de l'analyse des données massives pour nous assurer qu'une application particulière reposant sur des données massives peut être déployée en toute sécurité et être bénéfique pour l'ensemble de la société[14]. Les organisations doivent donc révéler la logique qui sous-tend l'analyse des données massives lorsqu'elle a un effet (direct ou indirect) sur l'individu. Elles doivent le faire de façon proactive[15], sans que les personnes concernées ne doivent prendre des mesures actives pour obtenir la divulgation[16].

Les données à caractère personnel traitées dans le cadre des données massives ne sont plus essentiellement composées d'informations que les personnes concernées ont sciemment données à des organisations. Une grande partie des données personnelles traitées est aujourd'hui observée ou déduite; en effet, l'enregistrement des activités en ligne et la localisation des smartphones et des tablettes ainsi que les possibilités accrues de suivre les activités dans le «monde réel» grâce à des dispositifs intelligents et à l'«Internet des objets» ajoutent à l'immense quantité de données à partir desquelles des déductions et des prédictions sont faites à notre sujet. La transparence est également importante lorsque des données sont collectées auprès de sources accessibles au public.

Que les données soient fournies volontairement, observées ou déduites[17] ou encore recueillies auprès de sources publiques, les personnes concernées ont parfaitement le droit de savoir quelles sont ces données, d'où elles viennent et auprès de qui les responsables du traitement les ont obtenues. Il devient de plus en plus nécessaire de fournir de façon plus proactive aux personnes concernées les données proprement dites «*sous une forme intelligible*» et de mentionner leur source.

La protection de la confidentialité commerciale ou des secrets d'affaires ne peut, de façon générale, primer sur les droits fondamentaux de la personne au respect de sa vie privée et à la protection de ses données. Au contraire, concilier les deux requiert une mise en balance soigneuse[18]. La décision de divulgation n'est pas non plus binaire. L'évaluation doit tenir compte des informations qui peuvent être divulguées ainsi que des procédures d'évaluation et de divulgation. Dans certains cas, il peut, par exemple, être fait appel à des tiers dignes de confiance, comme des assesseurs, plutôt que de dévoiler tous les détails à la personne concernée ou au public[19].

Les autorités chargées de la protection des données (et d'autres régulateurs, comme les autorités chargées de la protection des consommateurs, de la concurrence ou les régulateurs dans les domaines de la finance et de l'assurance) devraient également être en mesure de regarder dans la «boîte noire».

Pour les raisons exposées ci-dessus et afin de mettre un terme au profilage clandestin, nous recommandons que les dispositions relatives à la transparence dans la proposition de règlement européen sur la protection des données soient renforcées et incluent spécifiquement la divulgation de la «logique sous-tendant le processus décisionnel», des données proprement dites et de leur source.

2.2 De meilleurs outils pour informer les personnes concernées

Il est également important que des progrès soient réalisés dans la manière dont les informations sont divulguées aux personnes concernées. Toute information relative au traitement de données à caractère personnel doit utiliser un langage clair et simple, adapté au public visé, permettant aux personnes concernées de comprendre des informations complexes, et doit être aisément accessible. Si le traitement devient plus complexe, il incombe aux responsables du traitement de faire en sorte que les utilisateurs et les consommateurs soient mieux informés.

Ces politiques doivent véritablement servir à protéger les intérêts de la personne concernée par le traitement de données à caractère personnel et pas simplement à dégager le responsable du traitement de toute responsabilité légale. À l'instar de la législation relative à la protection des consommateurs, en cas d'ambiguïté dans ces politiques, elles doivent être interprétées en faveur de la personne concernée et non du responsable du traitement. Elles doivent aussi être honnêtes et sincères.

Des politiques plus traditionnelles de respect de la vie privée, tout en étant importantes pour la reddition de comptes, ne devraient pas être la seule, ni même la principale, source d'information des personnes concernées. Les autorités chargées de la protection des données recommandent depuis longtemps une notification «en couches»[20], qui informe les personnes concernées du traitement de leurs données étape par étape. Cela signifie que la personne concernée reçoit les informations essentielles relatives au traitement au moment où elle doit prendre une décision sur la base de ces informations (par exemple, une personne doit savoir si une application téléchargée pourra accéder à ses données de localisation avant de choisir de

l'installer) et d'autres informations dans d'autres formats, par exemple, par l'intermédiaire de renseignements plus détaillés sur un site Internet.

3. Au-delà de politiques illisibles sur le respect de la vie privée: le contrôle par l'utilisateur et le partage des avantages des données massives avec les personnes concernées

Au niveau mondial, le débat actuel est pollué par des malentendus sur les concepts de notification et de consentement. En droit européen de la protection des données[21], le consentement n'a jamais été synonyme de politiques longues et impénétrables sur le respect de la vie privée, rédigées par des juristes pour des juristes, auxquelles les utilisateurs doivent «consentir», à moins qu'ils ne veuillent renoncer totalement à l'utilisation du service souhaité. À l'inverse, le consentement est un choix véritable, librement consenti et assorti de la possibilité de dire «oui» sans le moindre préjudice. Il requiert également une compréhension claire de ce à quoi l'on souscrit.

Le consentement n'est pas toujours requis pour le traitement de données par des organisations[22]. Cependant, lorsqu'un consentement est requis, il doit être véritable: le simple fait de cocher une case sans comprendre à quoi vous consentez et sans véritable choix de le faire ou non, ne suffit pas à donner son consentement à des applications complexes reposant sur des données massives. La transparence et le contrôle par l'utilisateur doivent devenir une réalité[23].

3.1 Défaut de consentement: le droit d'opposition et les mécanismes de retrait

Le droit de s'opposer au traitement (qui n'est pas fréquemment exercé dans la pratique actuellement) peut se transformer en outil puissant dans les mains des parties prenantes lorsqu'il est appliqué comme un retrait inconditionnel «sans poser de questions». Dans certaines circonstances, cela peut contribuer à trouver le juste milieu entre le droit de la personne concernée à avoir un certain contrôle sur ses données et la flexibilité requise pour que les entreprises se développent, et innovent et exploitent au mieux l'immense quantité de données générées en ligne et hors ligne[24].

Un retrait inconditionnel signifie qu'une personne concernée est consciente que ses données sont traitées et qu'elle sait qu'elle peut s'y opposer si elle le décide. Elle peut ou non admettre totalement le fait que ses données soient traitées, mais souvent elle n'est pas suffisamment affectée négativement pour modifier les paramètres par défaut – ou ne s'en donne simplement «pas la peine». Le mécanisme de retrait influence subtilement la personne concernée à accepter le traitement, sans toutefois lui dénier totalement le droit de ne pas être d'accord.

Dans les cas limites, notamment, lorsque la mise en balance de l'intérêt légitime du responsable du traitement et les droits et intérêts des personnes concernées est difficile à atteindre, un mécanisme viable et bien conçu de retrait, tout en ne donnant pas nécessairement aux personnes concernées tous les éléments qui constitueraient un consentement valable en vertu de la législation européenne relative à la protection des données[25], pourrait grandement contribuer à préserver les droits et intérêts des personnes concernées[26].

En tant que société, nous devons choisir avec sagesse les conditions dans lesquelles nous exigeons des responsables du traitement qu'ils obtiennent un consentement véritable et quand nous devons nous contenter d'une simple évaluation de la mise en balance des intérêts et d'un retrait. Nous devons tout particulièrement nous efforcer de distinguer le traitement des données dont les avantages sont généraux ou sociétaux de ceux qui n'apportent que des avantages économiques aux organisations qui traitent les données. Nous devons également évaluer l'impact potentiel sur les personnes concernées et mettre soigneusement les deux en balance, tout en tenant compte d'autres facteurs pertinents[27].

Le retrait peut être facilité par des accords sectoriels, pour autant qu'ils soient efficaces et aisés à appliquer. Cependant, davantage d'efforts doivent être consentis avant qu'une initiative donnée puisse être approuvée, étant donné que jusqu'à présent, l'expérience nous a appris que ces accords ont donné peu de résultats concrets[28].

3.2 Au-delà du consentement: le contrôle des utilisateurs et le partage des avantages

Droit d'accès et portabilité des données

Le droit d'accès et de rectification des données à caractère personnel est l'un des principes fondamentaux de la législation européenne relative à la protection des données[29] et revêt une importance croissante en raison des avancées en matière d'analyse des données massives. Les personnes concernées doivent être formées pour mieux déceler les préjugés injustes et contester les erreurs dues à la logique algorithmique destinée à élaborer les hypothèses et les prévisions; le renforcement du droit d'accès et de rectification est une condition préalable à ceci.

Néanmoins, quelques rares personnes exercent ces droits dans la pratique[30]. L'une des raisons pour lesquelles ces droits d'accès potentiellement importants ne se sont pas transformés en outils plus puissants dans la pratique est le fait que souvent, les personnes concernées n'ont pas le temps ou ne voient pas l'intérêt de «*se battre pour la transparence et l'accès pour leur propre bien*»[31]. Toutefois, si les personnes concernées disposaient de la possibilité d'utiliser leurs données à caractère personnel pour en tirer concrètement parti, la situation pourrait changer. Cela pourrait se faire en «caractérisant» la protection des données: plutôt qu'une charge administrative, la fourniture de droits d'accès pourrait devenir une caractéristique du service offert aux clients[32]. Un exemple tiré du quotidien est l'accès aux informations bancaires en ligne.

Alors que les données massives se développent, les organisations utilisent des données à caractère personnel pour des finalités secondaires qui ne sont pas strictement nécessaires à la fourniture des services. Si elles souhaitent le faire, elles devraient être prêtes à partager les profits générés par le traitement des données à caractère personnel avec les personnes concernées dont les données sont traitées[33]. Il s'agit d'une exigence d'équité fondamentale; elle n'est toutefois par un simple impératif éthique.

Les données sont souvent comparées à d'autres ressources, comme le pétrole, qui sont commercialisées idéalement entre des parties à la transaction disposant du même degré élevé d'information. Les marchés des données à caractère personnel sont, toutefois, loin d'être transparents, équitables ou efficients. Les clients ignorent généralement la valeur exacte des données à caractère personnel qu'ils fournissent en échange de «services gratuits». De ce fait, ils ne sont pas correctement dédommagés pour les renseignements personnels qu'ils communiquent.

Comment et dans quelle mesure les personnes concernées devraient-elles bénéficier des avantages tirés du traitement de leurs données personnelles sont deux questions essentielles qui devraient faire l'objet d'une réflexion dans le cadre de la réalisation du marché unique numérique.

Une des façons d'accroître le contrôle par les personnes concernées, de partager avec elles les avantages tirés des données massives et, dans le même temps, de créer des incitations pour un traitement efficient et transparent des données à caractère personnel, passe par la portabilité des données. La portabilité des données impliquerait que les organisations:

- donnent aux personnes concernées un accès à leurs propres données dans un format portable, interopérable et lisible par une machine (en d'autres termes, utilisable et réutilisable);
- leur permettent de modifier, supprimer, transférer ou traiter ultérieurement les données les concernant;
- leur permettent de changer de fournisseurs (par exemple, transférer leurs photos, leurs archives bancaires ou de santé ou leurs courriels à un autre prestataire de services) et
- leur permettent de profiter des applications d'autres tiers pour analyser leurs propres données et en tirer des conclusions utiles (par exemple, changer d'habitudes alimentaires ou sportives, rechercher des soins personnalisés, prendre des décisions d'investissement plus avisées, opter pour un fournisseur d'électricité moins cher).

Autoriser la portabilité des données pourrait permettre aux entreprises et aux personnes concernées de maximaliser les avantages des données massives de manière plus équilibrée et transparente et pourrait contribuer à résorber le déséquilibre économique entre les responsables du traitement, d'une part, et les personnes concernées, de l'autre. La portabilité des données pourrait aussi permettre aux personnes concernées de tirer profit de la valeur produite par l'utilisation de leurs données à caractère personnel: elles pourraient utiliser les données à des fins qui leur sont propres ou accorder une licence pour l'utilisation ultérieure par des tiers, en échange de services supplémentaires ou d'une rémunération en espèce. En outre, la portabilité pourrait également contribuer à réduire les pratiques déloyales ou discriminatoires et les risques d'utilisation de données inexactes à des fins décisionnelles.

De plus, la portabilité des données est non seulement une bonne chose pour la protection de celles-ci, mais aussi pour la concurrence et la protection des consommateurs. Elle peut notamment favoriser un environnement de marché plus compétitif en permettant aux clients de changer plus aisément de fournisseurs (par exemple, dans le domaine de la banque par Internet ou dans le cas de fournisseurs d'énergie dans un réseau intelligent). Par ailleurs, elle peut également contribuer au développement d'autres services à valeur ajoutée par des tiers, qui peuvent accéder aux données des clients à la demande et sous réserve du consentement de ces derniers. Une fois encore, cela peut réduire les obstacles à l'entrée sur de nouveaux marchés qui ont besoin d'accéder à des données à caractère personnel et contribuer à créer des structures de marché plus concurrentielles et moins monopolistiques[34].

Compte tenu de ses avantages, le CEPD soutient fermement l'inclusion d'un droit à la portabilité des données dans la proposition de règlement européen sur la protection des données ainsi que l'inclusion de ce droit dans la législation, la réglementation ou les orientations sectorielles, le cas échéant (par exemple, en ce qui concerne les compteurs

intelligents). Nous encourageons également les initiatives gouvernementales ou privées visant à faciliter la portabilité des données.

Espaces de données personnelles

En complément de la portabilité des données et prenant appui sur celle-ci, une méthode visant à améliorer le contrôle des personnes concernées sur leurs données et déterminant qui peut y accéder et à quelles fins, pourrait être l'utilisation d'espaces de données personnelles (également appelés «entrepôt de données» ou «coffres de données»). Parmi les exemples de «données massives» mises à jour en temps réel qui peuvent être stockées dans un espace de données personnelles, citons la localisation d'une personne suivie par des capteurs dans sa voiture ou dans son téléphone portable, la tension artérielle ou d'autres données relatives à la santé ou à la forme physique suivies par un appareil médical ou un traceur de fitness.

La communication de la Commission européenne sur les données massives[35] fait spécifiquement référence à l'utilisation d'«espaces de données personnelles»[36], en tant que lieux sûrs et sécurisés, axés sur l'utilisateur, pour le stockage et la commercialisation éventuelle de données à caractère personnel. Nous sommes également d'avis que des outils numériques et des modèles commerciaux innovants, reposant sur l'autonomisation des consommateurs, devraient être encouragés. Ces outils incluent des mécanismes permettant aux personnes concernées de participer à l'utilisation et à la diffusion de leurs données et de tirer profit de ce partage de données.

Ceci pourrait potentiellement permettre de passer de modèles commerciaux dans lesquels les organisations suivent de plus en plus le comportement en ligne et hors ligne des personnes sans que ces dernières en soient pleinement informées et y aient consenti à un modèle dans lequel les personnes concernées gèrent leurs propres informations à des fins qui leur sont propres et partagent certaines de ces informations lorsqu'elles le souhaitent, avec qui elles le veulent, à une valeur juste et sous réserve de protections adéquates[37]. Les entrepôts de données personnelles pourraient aider à dissiper certaines des inquiétudes concernant la perte du contrôle individuel sur les données à caractère personnel, qui a été mentionnée plus haut comme étant l'une des préoccupations principales soulevées par les données massives[38].

L'Union européenne devrait se pencher sur la question de savoir comment promouvoir des outils et des produits interopérables, conviviaux, dignes de confiance et fiables et examiner les avantages et les contraintes ainsi que les défis technologiques qui y sont liés[39].

3.3 De nouvelles manières innovantes de fournir des informations, un accès et un contrôle aux personnes concernées

Les entreprises et d'autres organisations qui déploient de gros efforts pour trouver des manières innovantes d'utiliser les données à caractère personnel devraient faire preuve du même esprit innovant lorsqu'elles conçoivent de **nouvelles manières de fournir des informations, un accès et un contrôle aux personnes concernées**[40].

De nouvelles méthodes conviviales pour l'utilisateur devraient être développées et proposées afin de permettre aux personnes concernées de donner ou de refuser leur consentement de manière éclairée. L'octroi aux personnes concernées d'un certain degré de contrôle sur l'utilisation qui est faite de leurs données est souvent une obligation légale ou une bonne pratique qui aide également les responsables du traitement à instaurer un climat de confiance.

Ainsi, les personnes concernées devraient pouvoir activer et désactiver sans effort le suivi ou le partage d'informations sur les appareils et applications qu'elles utilisent, en fonction de la localisation, de l'heure et de la date, aussi bien par application que de façon globale. Il conviendrait aussi de proposer de meilleures solutions pour faciliter la rectification, la mise à jour ou la suppression de données ou pour modifier les personnes qui peuvent y avoir accès, contrôler qui a effectivement accédé aux données et pour quelles finalités. Cela nous amène à notre sujet suivant, la protection des données et le respect de la vie privée dès la conception.

4. Protection des données et respect de la vie privée dès la conception

Le concept de respect de la vie privée et de protection des données dès la conception a pour but d'intégrer le respect de la vie privée et la protection des données dans les spécifications de conception et l'architecture des systèmes et des technologies d'information et de communication. Il ne se limite pas aux aspects techniques et les mesures organisationnelles sont tout aussi importantes.

Une technologie et une ingénierie respectueuses de la vie privée peuvent jouer un rôle essentiel pour faire de la transparence et du contrôle par l'utilisateur, tel que décrit plus haut, une réalité. La législation, la réglementation, les conditions contractuelles, les procédures internes et les politiques relatives au respect de la vie privée, si importantes soient-elles, ne suffiront pas à elles seules. **De nouvelles manières innovantes d'être informées de ce qui est fait avec leurs données et d'exercer un contrôle sur celles-ci** doivent être proposées aux personnes concernées. Cela nécessite une ingénierie innovante et respectueuse de la vie privée ainsi que des mesures organisationnelles et des pratiques commerciales respectueuses de la vie privée. Une ingénierie innovante et responsable peut, notamment, faciliter l'exercice par les personnes concernées de leur droit d'accès, d'opposition, de retrait, de rectification ainsi que de leur droit à la portabilité des données. Une ingénierie respectueuse de la vie privée peut également être précieuse en contribuant à développer de nouveaux modèles commerciaux pour produire de la valeur à partir, par exemple, des entrepôts de données.

Un autre domaine totalement différent, dans lequel des solutions d'ingénierie innovante doivent être trouvées, est lié au concept de la «**séparation fonctionnelle**». Lorsqu'une organisation qui traite des données veut uniquement détecter les tendances et les corrélations dans les données plutôt que d'appliquer directement la vision d'ensemble qu'elle s'est faite sur les personnes concernées, la «séparation fonctionnelle» pourrait contribuer à réduire les effets sur les droits des personnes, tout en permettant aux organisations d'exploiter des utilisations secondaires des données[41]. La séparation fonctionnelle vise à prendre des mesures techniques et organisationnelles afin que les données utilisées à des fins de recherche ne puissent pas être utilisées ultérieurement pour «étayer des mesures ou des décisions» en rapport avec les personnes concernées (sauf si ces dernières y ont expressément consenti)[42].

Par ailleurs, et en dépit de leurs limites, des **techniques d'anonymisation** adéquates peuvent toujours contribuer à garantir une utilisation ou un partage sûr des données au sein d'une organisation, entre des organisations différentes ou lorsque les données sont rendues publiques, comme dans le cas des projets de «données ouvertes»[43]. Pour rendre des données anonymes, il ne suffit pas de supprimer des attributs directs d'identité d'un ensemble de données. Plus un ensemble de données devient volumineux et exhaustif, plus il existe de possibilités d'identifier les personnes concernées auxquelles les données se rapportent, en

particulier lorsque les données sont conservées pendant des périodes plus longues et/ou lorsqu'elles sont partagées[44]. Une utilisation prudente de ces techniques, alliée à d'autres mesures de protection (comme des restrictions concernant la durée de rétention des données, le contrôle d'accès), peut toutefois aider, dans certains cas, à assurer le respect de la législation relative à la protection des données.

Enfin, les initiatives et les investissements en faveur de l'utilisation et du déploiement des données massives doivent considérer une sécurité adéquate comme une condition préalable essentielle à une utilisation socialement acceptable des données massives et les mesures de sécurité et d'évaluation des risques comme faisant partie intégrante des données massives.

5. Reddition de comptes

La reddition de comptes impose aux responsables du traitement d'instaurer des mécanismes et des systèmes de contrôle internes qui assurent la conformité et fournissent des preuves – notamment des politiques internes et des rapports d'audit – démontrant la conformité aux parties prenantes externes, y compris les autorités de contrôle. La reddition de comptes n'est pas un exercice ponctuel. Une vérification régulière du fait que les systèmes de contrôle interne sont toujours adaptés et que chaque traitement de données est conforme à la législation est un élément essentiel de la reddition de comptes.

Un grand nombre d'éléments entrent dans la composition de bonnes pratiques responsables. Le respect de la vie privée et la protection des données dès la conception et par défaut, les évaluations d'impact de la protection des données, les audits et la certification ainsi que la disponibilité d'experts compétents en matière de protection des données, y compris un délégué à la protection des données, au sein de l'organisation peuvent contribuer à un système de contrôle interne responsable et en font partie intégrante; et ces éléments devraient être imposés et encouragés, selon les cas, car ils peuvent jouer un rôle important dans une utilisation responsable des données massives.

Dans le cas de l'analyse de données massives, il est souvent difficile de décider ce qui est juste et licite et ce qui ne l'est pas.

Voici quelques exemples de décisions essentielles qu'une organisation responsable doit prendre en vertu de la législation européenne relative à la protection des données:

- l'utilisation secondaire des données est-elle conforme au principe de la limitation de la finalité;
- les données initialement utilisées dans un contexte peuvent-elles être considérées comme adéquates, pertinentes et proportionnelles pour être réutilisées dans un autre contexte et
- en l'absence du consentement des personnes concernées, une organisation peut-elle invoquer son intérêt légitime au traitement de données.

Bien que ces évaluations reposent sur des exigences légales, elles requièrent souvent un exercice global de mise en balance et la prise en compte de multiples facteurs, notamment celui de savoir si le traitement des données répond aux attentes légitimes des personnes concernées, s'il est susceptible d'aboutir à une discrimination injuste ou s'il peut avoir un

autre effet négatif sur les personnes concernées ou sur la société dans son ensemble. Ces évaluations soulèvent souvent des questions complexes d'éthique et d'équité sur le plan commercial et ne sauraient être réduites à un simple exercice mécanique consistant à cocher des cases sur la conformité. Plus les ordinateurs deviennent puissants, plus le défi est difficile à relever. Ainsi, la recherche a conclu que les ordinateurs sont plus précis que les hommes pour tirer des prédictions en se fondant sur des «empreintes digitales», telles que des traits de caractère, des attitudes politiques et la santé physique[45].

C'est la raison pour laquelle ces évaluations devraient être réalisées par une équipe pluridisciplinaire (par exemple, des informaticiens, des ingénieurs, des juristes, des délégués à la protection des données, des statisticiens, des spécialistes des données, des médecins, des scientifiques et des experts en marketing, en assurance ou en finance).

Des « comités d'éthique» peuvent, le cas échéant, jouer un rôle dans l'élaboration de procédures internes plus responsables. À la manière d'organes similaires dans le domaine de la recherche scientifique, ils pourraient formuler des recommandations ou adopter des décisions contraignantes au sein de l'organisation, sur la possibilité légale et éthique des types particuliers d'analyse de données massives. D'autres dispositions organisationnelles peuvent toutefois se révéler aussi efficaces. L'important est d'instaurer un cadre de mise en conformité qui contribuera à faire ce que les décisions qui seront finalement prises sur un traitement de données soient «éthiques», «justes» et «licites».

6. Prochaines étapes: traduire les principes dans la pratique

Pour relever les défis que posent les données massives, nous devons **permettre l'innovation tout en protégeant les droits fondamentaux**. Pour y parvenir, les principes établis dans la législation européenne relative à la protection des données devraient être préservés, mais appliqués de façon innovante.

6.1 Une réglementation tournée vers l'avenir

Les négociations relatives à la proposition de règlement général sur la protection des données sont entrées dans leur phase finale. Nous avons instamment prié les législateurs européens d'adopter un paquet de réforme de la protection des données qui renforce et modernise le cadre réglementaire afin qu'il demeure efficace à l'ère des données massives, tout en développant la confiance des personnes concernées dans la sécurité en ligne et le marché unique numérique[46].

Dans l'avis n° 3/2015, qui est assorti de recommandations sur le texte complet du règlement proposé, nous avons clairement indiqué que nos principes actuels en matière de protection des données, notamment la nécessité, la proportionnalité, la minimisation des données, la limitation de la finalité et la transparence, doivent rester des principes clés. Ils doivent être la base de référence dont nous avons besoin pour protéger nos droits fondamentaux dans le monde des données massives[47].

Dans le même temps, ces principes doivent être renforcés et appliqués plus efficacement et de façon plus moderne, souple, créative et innovante. Ils doivent aussi être complétés par de nouveaux principes, comme la responsabilité et le respect de la vie privée et la protection des données dès la conception et par défaut.

Une transparence accrue, de puissants droits d'accès aux données et de portabilité de celles-ci, ainsi que des mécanismes efficaces de retrait peuvent constituer une condition préalable pour permettre aux utilisateurs de mieux contrôler leurs données et peuvent également contribuer à des marchés plus efficaces pour les données personnelles, dans l'intérêt des consommateurs et des entreprises.

Enfin, élargir le champ d'application de la législation européenne sur la protection des données aux organisations qui ciblent des personnes dans l'UE et doter les autorités chargées de la protection des données de pouvoirs leur permettant de prendre des mesures correctrices utiles, y compris des amendes efficaces, comme le prévoit la proposition de règlement, seront des exigences essentielles pour une application efficace de notre législation sur le plan international. Le processus de réforme joue un rôle crucial à cet égard.

Afin de garantir la mise en œuvre effective des règles, des autorités de protection des données indépendantes doivent être dotées non seulement de pouvoirs légaux et d'instruments solides, mais également des ressources nécessaires pour développer leur capacité, au rythme de la croissance des activités reposant sur les données.

6.2 Comment le CEPD fera-t-il avancer ce débat?

Une bonne réglementation, même si elle est essentielle, ne suffit pas. Les entreprises et les autres organisations qui déploient d'importants efforts dans la recherche de solutions innovantes pour l'utilisation des données à caractère personnel devraient faire preuve du même esprit innovant dans la mise en œuvre des principes de protection des données. Les autorités chargées de la protection des données devraient, elles aussi, faire appliquer la réglementation, récompenser le respect des règles et éviter d'imposer une charge administrative et une paperasserie inutiles.

Comme annoncé dans sa stratégie pour la période 2015-2019, le CEPD entend contribuer à favoriser ces efforts.

Nous avons l'intention de mettre sur pied un groupe consultatif externe sur la dimension éthique, composé de personnalités éminentes et indépendantes, possédant ensemble de l'expérience dans de multiples domaines «*afin d'explorer les relations entre les droits de l'homme, la technologie, les marchés et les modèles commerciaux au cours du XXI^{e} siècle*», d'analyser en profondeur l'impact des données massives, d'évaluer les changements qui en résultent pour nos sociétés et d'aider à identifier les questions qui devraient faire l'objet d'un processus politique[48].

Nous allons également élaborer un modèle de politiques honnêtes en matière d'information pour les organes de l'UE qui proposent des services en ligne, en vue de contribuer au développement de meilleures pratiques pour l'ensemble des responsables de traitement.

Enfin, nous allons favoriser les discussions, par exemple pour recenser, encourager et promouvoir les meilleures pratiques destinées à accroître la transparence et le contrôle par l'utilisateur et à étudier les possibilités des entrepôts de données personnelles et la portabilité des données. Le CEPD a l'intention d'organiser un atelier sur la protection des données massives à l'intention des décideurs et des personnes qui traitent de gros volumes de données personnelles dans les institutions de l'UE, ainsi que des experts externes, de déterminer si des orientations supplémentaires sont nécessaires et de faciliter le travail du réseau d'ingénierie de la vie privée sur l'internet (IPEN) en tant que centre de connaissances interdisciplinaires pour les ingénieurs et les spécialistes de la vie privée.

(signé)

Bruxelles, le 19 novembre 2015

Giovanni BUTTARELLI

Contrôleur européen de la protection des données

Notes

[1] Avis nº 4/2015 du CEPD.

[2] *Public Utilities Commission v. Pollak*, 343 U.S. 451, 467 (1952) (Juge William O. Douglas, juge dissident).

[3] Le 25 janvier 2012, la Commission européenne a adopté un train de mesures visant à réformer le cadre européen de la protection des données. Ces mesures comprennent: (i) une «communication» (COM(2012)9 final), (ii) une proposition de «règlement général sur la protection des données» («proposition de règlement») (COM(2012)11 final) et (iii) une proposition de «directive» sur la protection des données en matière pénale (COM(2012)10 final).

[4] «Les données massives renvoient à la croissance exponentielle de la disponibilité et de l'utilisation automatisée d'informations: elles renvoient à de gigantesques ensembles de données numériques détenus par les sociétés, les gouvernements et d'autres grandes organisations, qui sont ensuite analysées de manière approfondie (d'où le nom "analyse") en utilisant des algorithmes informatiques»; avis nº 3/2013 du groupe de travail «Article 29» sur la limitation de la finalité.

[5] Voici quelques exemples d'applications reposant sur des données massives qui se servent de données à caractère personnel (ces exemples montrent de quoi une technologie est capable et pas nécessairement ce qui est éthique ou légal):

- Recherche médicale et médecine personnalisée. Si des scientifiques reçoivent un accès à nos profils génétiques, à nos antécédents médicaux et à des données sur notre style de vie (par exemple, via des applications mobiles de santé et de forme physique, des informations des réseaux sociaux, des données sur des cartes de crédit et de fidélité), procéder à une analyse des données massives sur ces immenses séries de données précieuses pourrait potentiellement révolutionner la recherche médicale en permettant à des scientifiques de trouver de nouvelles corrélations et, au bout du compte, peut-être de nouveaux traitements pour des maladies. L'analyse des données massives pourrait également prédire si un patient est susceptible de contracter une maladie, est sujet à un effet secondaire ou réagit à certains types de traitements médicaux. Cela pourrait ensuite permettre aux médecins de prodiguer un traitement médical personnalisé et, donc, plus efficace.
- Des moteurs de recherche reposent sur des données massives et c'est également le cas de nombreux autres services en ligne relatifs à l'évaluation ou à la recommandation d'un contenu, de produits ou de services. La publicité comportementale et ciblée, les offres et les remises personnalisées ainsi que les recommandations personnalisées de contenu média, d'hôtels ou de restaurants sont des services qui reposent sur des données massives.
- La notation de la solvabilité exploite des données massives pour évaluer les risques de ne pas honorer nos dettes.
- Lutte contre la fraude fiscale: en autorisant les autorités fiscales à accéder à certaines données d'autres agences gouvernementales ou d'entreprises privées, elles pourraient, grâce aux données massives, croiser les bases de données fiscales avec d'autres informations, comme l'immatriculation des véhicules, les informations relatives aux cartes de crédit ou les informations détenues par des intermédiaires financiers pour trouver les personnes dont les schémas de dépenses/d'investissement et les impôts ne correspondent pas.
- Lutte contre le terrorisme et le crime organisé: grâce à l'interception des données de communication d'un grand nombre de personnes (suspectes ou non) et à leur criblage au moyen d'un outil d'analyse puissant, les agences de renseignement espèrent découvrir des attentats terroristes en préparation.

[6] Voir, par exemple :

- la résolution sur les données massives adoptée en octobre 2014 par la 36e conférence internationale sur la protection des données et les commissaires chargés du respect de la vie privée («résolution de la conférence internationale sur les données massives»;

- le Groupe de travail international sur la protection des données dans les télécommunications, document de travail sur les données massives et le respect de la vie privée (55e réunion, 5 et 6 mai 2014, Skopje) («document de travail du groupe de Berlin sur les données massives»);
- la déclaration du groupe de travail «Article 29» sur l'impact du développement des données massives sur la protection des personnes physiques à l'égard du traitement de leurs données à caractère personnel dans l'UE, adoptée le 16 septembre 2014 (WP 221) («déclaration du groupe de travail "Article 29" sur les données massives»);
- bureau de l'Information Commissioner's Office britannique, Guide des données massives et de la protection des données, juillet 2014 («Guide ICO sur les données massives»);
- rapport de l'autorité norvégienne de la protection des données «données massives et principes du respect de la vie privée sous pression», 2013 («rapport norvégien sur les données massives»)

[7] Voir «Big Data: Seizing Opportunities, Preserving Values», Bureau exécutif du Président, mai 2014, page 10.

[8] Avis 3/2013 du groupe de travail «Article 29» sur la limitation de la finalité, annexe 2.

[9] Les utilisations abusives de ce déséquilibre de pouvoir peuvent revêtir des formes diverses. La discrimination par les prix en est une. Elle permet aux entreprises de proposer des produits et des services à des prix différents à des personnes différentes dans une tentative de retirer le prix le plus élevé que chaque consommateur est prêt à débourser. De grands ensembles de données sur les comportements individuels sont déjà aisément accessibles et contiennent des informations potentiellement utiles pour une tarification adaptée aux différents publics. Le ciblage de clients vulnérables est une autre forme courante d'abus.

[10] Bruce Schneier, *Data and Goliath*, 2015, p. 238.

[11] Rapport norvégien sur les données massives, page 7, point 8.

[12] De plus, les données massives peuvent créer des «bulles de filtrage» (ou des «chambres d'écho» personnelles) pour les particuliers. Dans notre monde de plus en plus individualisé, les algorithmes tentent de deviner quelles informations chacun de nous voudrait voir fondées sur ce que l'on sait de nous (par exemple, notre localisation, notre navigation passée sur la Toile et notre historique de recherche et d'achat). Toutes les informations que nous recevons sont filtrées de façon de plus en plus opaque et complexe. Le danger existe qu'il devienne moins probable que nous trouvions des informations qui remettent en cause notre point de vue existant. Nous serons effectivement de plus en plus isolés dans nos propres bulles culturelles et idéologiques et séparés du reste de la société.

[13] Aux États-Unis, le Centre national sur la législation relative à la protection des consommateurs a constaté que les produits de crédit vendus selon des procédés non traditionnels reposant sur des données exigeaient des taux annuels compris entre 134 % et 748 %; *Big Data: A Big Disappointment for Scoring Consumer Credit Risk*, mars 2014. En août 2015, un brevet américain a été acheté; il couvrait une technologie pour évaluer le risque de crédit des membres du réseau social connecté à la personne concernée afin de déterminer s'il fallait traiter ou rejeter une demande de prêt; «*Facebook patent: Your friends could help you get a loan – or not*» (04.08.2015) http://money.cnn.com/2015/08/04/technology/facebook-loan-patent/. En octobre 2015, l'autorité bancaire européenne, l'Autorité européenne des marchés financiers, et l'Autorité européenne des assurances et des pensions professionnelles ont lancé une enquête conjointe sur les risques et les avantages des données massives.

[14] En ce qui concerne le concept des «boîtes noires» et l'importance de la transparence, voir par exemple «The Black Box Society, The Secret Algorithm That Control Money and Information», Frank Pasquale (Harvard University Press, 2015).

[15] Voir également les articles 10 et 11 de la directive 95/46/CE. Voir aussi l'article 15.

[16] Parmi les exemples tirés du quotidien pour lesquels «la logique sous-jacente au processus décisionnel» devrait être dévoilée, on peut citer un système d'assurance automobile personnalisé (en

utilisant les données du capteur du véhicule pour juger les habitudes de conduite), des services d'évaluation de la capacité de crédit, un système de commercialisation et de tarification qui détermine quelle ristourne une personne recevra ou quel contenu média doit être recommandé à un individu.

[17] Les données déduites englobent également le profil de la personne concernée, comme l'évaluation de sa capacité de crédit ou le résultat d'une évaluation de son état de santé.

[18] Dans le cadre de l'évaluation, il convient également de considérer que le secret n'est pas la seule manière de protéger véritablement des produits et services innovants. En effet, de nombreux algorithmes véritablement innovants peuvent aussi être protégés par des droits de propriété intellectuelle plutôt qu'en invoquant des secrets d'affaires. Les brevets, par exemple, offrent un niveau élevé de protection de la propriété intellectuelle tout en assurant une plus grande transparence vis-à-vis de la personne concernée.

[19] Sur la «transparence qualifiée», voir, par exemple, Frank Pasquale: «The Black Box Society», p. 160-165.

[20] Voir, par exemple, avis 10/2004 du groupe de travail «Article 29» intitulé «Dispositions davantage harmonisées en matière d'informations» (WP 100), avis 2/2009 du groupe de travail «Article 29» sur la protection des données à caractère personnel de l'enfant (Principes généraux et cas particulier des écoles) (WP 160), avis 3/3013 du groupe de travail «Article 29» sur la limitation de la finalité, page 16, exemples 9, 10 et 11 à l'annexe 3, pages 52 et 53.

[21] Voir l'article 7, point a), de la directive 95/46/CE.

[22] L'avis 6/2014 du groupe de travail «Article 29» sur la notion d'intérêt légitime énonce des orientations et un ensemble de critères afin d'aider à déterminer dans quels cas une organisation peut invoquer l'intérêt légitime et dans quels cas elle doit obtenir le consentement des personnes concernées.

[23] Par ailleurs, la protection de la vie privée est souvent faussement assimilée au choix de savoir si des personnes reçoivent ou non une publicité ciblée en ligne et les «tableaux de bord» conviviaux donnent fréquemment l'illusion d'un contrôle, sans toutefois réellement permettre aux personnes concernées de refuser d'être suivies et de profiter malgré tout des avantages de l'Internet. Il n'y a pas que la publicité ciblée proprement dite qui peut entraîner une violation de la vie privée; il y a aussi l'incapacité des personnes concernées à éviter d'être suivies.

[24] Si une organisation souhaite invoquer l'article 7, point f), de la directive 95/46/CE (intérêt légitime) comme base juridique du traitement de données à caractère personnel, elle doit donner aux personnes concernées le droit d'opposition visé à l'article 14, point a), de la directive, sous certaines conditions. Outre le droit d'opposition consacré par l'article 14, point a), une organisation peut toutefois aussi décider de proposer un droit de retrait plus large, inconditionnel et généralement applicable aux personnes concernées. Dans le cas du marketing direct, ce retrait est déjà une obligation légale en vertu de l'article 14, point b), de la directive 95/46/CE. Il s'agit du type de retrait «sans poser de questions» évoqué dans la présente section. Voir également l'avis 6/2014 du groupe de travail «Article 29» sur l'intérêt légitime, pages 44 et suivantes, sous l'intitulé «*Le droit d'opposition et au-delà*».

[25] En vertu de l'article 7, point a), de la directive 95/46/CE.

[26] L'avis 6/2014 du groupe de travail «Article 29» sur l'intérêt légitime donne des orientations et des exemples, mais insiste sur l'importance d'une évaluation au cas par cas. Voir, en particulier, les pages 31 à 33 et 44 à 47, ainsi que les exemples 4 et 5 de l'annexe 2 à l'avis, page 59.

[27] Idem.

[28] De nombreuses critiques ont été émises, notamment, sur les initiatives sectorielles concernant le retrait de la publicité comportementale en ligne et du refus de suivi. Voir, par exemple, http://www.nytimes.com/2014/12/27 /opinion/the-slow-death-of-do-not-track.html?_r=0.

[29] Voir l'article 8, paragraphe 2, de la Charte des droits fondamentaux de l'Union européenne et l'article 12 de la directive 95/46/CE.

[30] Voir, par exemple, p. 26 et suivantes, Omer Tene et Jules Polonetsky (2012, «Big Data for All: Privacy and User Control in the Age of Analytics, 11 Northwestern Journal of Technology and Intellectual Property 239 (2013), disponible sur http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2149364.

[31] Idem.

[32] Idem.

[33] Cela s'ajoute à ce qui devrait être un impératif de bon sens et va au-delà, à savoir minimiser et compenser les effets externes négatifs que les entreprises de données massives peuvent causer, comme des risques accrus pour la sécurité des personnes dont les données sont traitées.

[34] Voir également le paragraphe 26 de l'avis préliminaire du CEPD sur la vie privée et la compétitivité à l'ère de la collecte de données massives, adopté le 26 mars 2014. En outre, les avantages de la portabilité des données ont également été mis en évidence par le groupe de travail «Article 29» dans son avis 3/2013 sur la limitation de la finalité et dans son avis 6/2014 sur l'intérêt légitime. Les deux avis font également tous deux spécifiquement référence à des initiatives comme les «midata» au Royaume-Uni (et son équivalent en France), qui s'appuient sur le principe de base voulant que les données devraient être «retransmises» aux consommateurs afin qu'ils puissent les utiliser à des fins qui leur sont propres. Pour en savoir plus sur les «midata» au Royaume-Uni et des initiatives similaires en France, voir:
http://www.midatalab.org.uk/,
https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/34747/12-983-midata-company-briefing-pack.pdf et http://mesinfos.fing.org/.

[35] Communication de la Commission «Vers une économie de la donnée prospère», qui présente la stratégie de la Commission en matière de données massives (COM(2014)442 final).

[36] Section 4.2.3.1, paragraphe 4.

[37] Voir, par exemple, Doc Searls, *The Intention Economy: When Customers Take Charge* (Boston: Harvard Business Review Press, 2012).

[38] Voir, par exemple, Ira S. Rubinstein, *Big Data: The End of Privacy or a New Beginning?* International Data Privacy Law, 2013, Vol 3, No 2.

[39] La sécurité, la fiabilité et la faisabilité technique sont quelques-uns des problèmes à résoudre. Il conviendrait également de clarifier à qui bénéficient les espaces de données personnelles. Enfin, il est important d'observer que les espaces de données personnelles ne devraient pas «vendre» des données à caractère personnel, mais plutôt permettre à des tiers d'«utiliser» des données à caractère personnel à des fins spécifiques et pendant des périodes spécifiques, sous réserve des conditions fixées par les personnes concernées elles-mêmes et de toutes les autres mesures de protection des données.

[40] Le consentement devrait être suffisamment pointu et – même si c'est la même entité qui procède au traitement ultérieur des données – il devrait couvrir les différentes activités de traitement de données pour chaque service. La combinaison de données pour une finalité différente nécessite aussi un consentement spécifique.

[41] Voir les pages 27, 29 et 30, et l'annexe 2, page 46, de l'avis 3/2013 du groupe de travail «Article 29» sur la limitation de la finalité.

[42] Il existe peu de preuves d'une expérience relative à une application réelle d'une séparation fonctionnelle en dehors de quelques organisations spécialisées, tels que les offices nationaux de statistiques et les instituts de recherche. Pour exploiter pleinement les utilisations secondaires des données, il est essentiel que d'autres organisations développent leur expertise et offrent des garanties comparables contre toute utilisation abusive des données.

[43] Pour en savoir plus sur la manière de déterminer quand des ensembles de données agrégées peuvent être publiés comme données ouvertes, voir la section 6 de l'avis 6/2013 du groupe de travail «Article 29» sur la réutilisation des informations du secteur public (ISP) et des données ouvertes.

[44] Une analyse des techniques d'anonymisation actuellement disponibles a été réalisée par le groupe de travail «Article 29» dans son avis 5/2014 sur les techniques d'anonymisation.

[45] Les évaluations informatisées de la personnalité sont plus précises que celles réalisées par des humains (Wu Youyoua, Michal Kosinski et David Stillwell). Décembre 2014.

[46] Avis nº 3/2015 du CEPD.

[47] Nous devons résister à la tentation d'atténuer le niveau actuel de protection pour essayer de répondre au besoin perçu d'une approche réglementaire plus laxiste en matière de données massives. La protection des données doit continuer de s'appliquer à l'ensemble du traitement, c'est-à-dire pas uniquement à l'utilisation des données, mais aussi à leur collecte. Rien ne justifie un blanc-seing pour le traitement de données pseudonymes ou de données accessibles au public. La définition des données à caractère personnel doit demeurer inchangée, mais le texte du règlement proprement dit pourrait la clarifier davantage. Elle doit, en effet, couvrir toutes les données concernant une personne qui est identifiée ou isolée ou peut être identifiée ou isolée, que ce soit par le responsable du traitement ou par un autre tiers.

[48] Avis nº 4/2015 du CEPD.