MÉGADONNÉES ET PROTECTION DE LA VIE PRIVÉE

par BERTIL COTTIER*

1. Introduction

La régulation de l'utilisation des mégadonnées (c'est ainsi que l'on désigne le *Big data* en bon français) est un nouveau défi pour la protection de la vie privée. Un de plus pour un domaine constamment malmené par le développement fulgurant des nouvelles technologies de traitement et de communication de l'information.

On sait que la vie privée avait déjà été mise à rude épreuve par l'avènement, au cours de la seconde moitié du siècle précédent, des premiers ordinateurs et de leurs capacités inédites de stocker et d'analyser des données personnelles. Pour parer aux abus de l'informatique (entre autres la surveillance des citoyens par l'Etat), des instruments juridiques topiques ont été mis en place: les fameuses lois sur la protection des données, telle, au niveau de l'Union européenne, la directive 95/46 CE[1]. Des instruments certes innovateurs, mais qui se sont avérés vite dépassés par le progrès technique, à commencer par Internet à qui l'on doit la mondialisation de la communication, la permanence des informations mises en ligne et leur aisée repérabilité par le biais des moteurs de recherche. Autant de nouvelles menaces sur la vie privée qui ont imposé une révision des normes protectrices. Après bien des années d'hésitations et de controverses, cette révision est à bout touchant: fin avril 2016, l'Union Européenne adoptait un règlement général sur la protection des données[2]; un texte d'application directe[3] qui, passé un délai de mise en conformité de deux ans, supplantera ladite directive et les législations nationales des Etats membres qui l'implémentent.

Un nouveau cadre européen de la protection des données va donc voir le jour. Cela dit, ce cadre, même s'il est sensé améliorer les possibilités pour l'individu de contrôler l'utilisation de ses données par des tiers (qu'il s'agisse des autorités publiques ou d'organisations et d'entre

* *Professeur de droit de la communication – Faculté des sciences de la communication de l'Université de la Suisse italienne – via Buffi 13 – 6900 Lugano, e-mail: bertil.cottier@usi.ch*

prises privées) risque d'être obsolète à peine entré en vigueur. Il ne tient en effet pas compte de la véritable révolution qu'est le traitement des mégadonnées, un tournant radical qui affecte non seulement la quantité (phénoménale[4]) de données qui peuvent être exploitées, mais aussi la qualité des résultats obtenus: le profilage des individus devient beaucoup plus étroit, les contours de leurs activités, leurs modes de vie et leurs préférences, passés et présents, apparaissent toujours plus nettement. En d'autres termes, d'une image de faible résolution, on est passé à un portrait en «ultra-haute définition«.

La présente contribution entend mettre en lumière les menaces sur la vie privée posées par le traitement des mégadonnées et passer en revue les remèdes envisageables. Faute de place, on devra se contenter d'esquisser des solutions possibles. Cela dit, il y a lieu de souligner d'emblée que celles-ci se fondent sur une approche conciliatrice: jamais il n'a été question de bannir le *Big data*. Même les plus fervents défenseurs de la vie privée reconnaissent que les mégadonnées contribuent fortement au bien-être des citoyens, améliorant notamment les performances de la médecine, de la lutte contre la criminalité ou encore des services sociaux[5]. Il ne s'agit donc pas d'interdire purement et simplement, mais de contrecarrer les dérapages dommageables à la vie privée[6].

2. Le régime de la protection des données en bref

Il est impossible de réaliser combien le traitement des mégadonnées a bouleversé le régime classique de la protection des données sans en rappeler d'abord les grandes lignes.

En Europe[7], les législations sur la protection des données régissent de façon circonstanciée toute opération (le terme consacré est *traitement*) relative à des données personnelles, notamment leur collecte, leur conservation, leur modification, leur transmission et leur destruction. Le traitement de ces données est soumis au respect d'une poignée de principes directeurs qui tendent à sauvegarder la vie privée des personnes dont les données sont traitées (*personnes concernées*) et à leur assurer une certaine maîtrise sur le sort de leurs données. Ainsi les personnes concernées doivent avoir connaissance de la nature du traitement et de sa finalité (principe dit de transparence); en outre, le traitement ne peut porter que sur les seules informations nécessaires au regard du but poursuivi (principe dit de *pertinence et minimisation* des données); de surcroît, les données ne doivent pas être utilisées pour un autre objectif que celui qui a justifié leur collecte (principe dit de *finalité*); enfin les données doivent être sécurisées contre l'interception par des tiers non autorisés. Les lois sur la protection des données octroient également à la personne concernée le privilège de s'opposer à un traitement qui lui serait contraire (*droit de blocage*) et celui de savoir qui traite des données à son sujet et dans quel but (*droit d'accès*). Enfin, elles instituent des organismes de contrôle indépendants, appelés à promouvoir et à mettre en

œuvre la protection de données; destinés à superviser le secteur privé et le secteur public, ces organismes doivent être indépendants de l'Etat.

Essentiels, ces principes de base sont au cœur des premiers instruments internationaux de protection des données (la directive 95/46 CE, mais aussi l'article 8 de la Charte européenne des droits fondamentaux de 1999[8] et la convention 108 du Conseil de l'Europe[9]) ; ils ont été ensuite confirmés dans les textes destinés à moderniser ces instruments, respectivement le règlement général (UE) 2016/679[10] et le projet de protocole d'amendement de la convention 108 (septembre 2016[11]).

3. Les nouvelles menaces sur la vie privée

3.1. *Le manque de transparence*

L'opacité du traitement est certainement l'un des aspects les plus négatifs du *Big data*, car elle en affecte tous les stades, de la collecte des données au cercle de diffusion des résultats en passant par la finalité de l'analyse. Une opacité souvent renforcée par une chaîne de traitement complexe, mettant en jeu une pluralité de sous-traitants dont certains sont situés à l'étranger (complexité encore accentuée par le recours fréquent à l'informatique en nuage – cloud – et l'ignorance du lieu de stockage des données qu'elle peut engendrer). Mais voyons cela de plus près.

Jusqu'à peu, l'exploitation des données personnelles nécessitait la création d'amas d'informations structurés, le plus souvent des fichiers électroniques ordonnant les entrées suivant des rubriques précises. L'organisation des données était un gage d'efficacité du traitement. Le *Big data* en revanche se passe de tout agencement logique des données: les puissants logiciels de traitement des mégadonnées permettent, à la vitesse de l'éclair, de combiner et d'analyser des données diverses de par leur nature (photos, vidéos, sons, graphiques, *likes*, métadonnées[12], etc.) et de par leur source (réseaux sociaux, moteurs de recherche, courriers électroniques, blogs, forums de discussions, etc.). Conséquence de cette hétérogénéité, le citoyen ignore quelles données le concernant sont utilisées et à quelles fins. Ce d'autant que la collecte des données elle-même n'est guère plus visible; le plus souvent l'agrégation des vastes volumes de données nécessaires à l'analyse, se fait à l'insu de la personne concernée.

Reste que l'on doit admettre que la personne concernée n'est pas toujours exempte de reproche. Ne contribue-t-elle pas, plus ou moins involontairement, à l'opacité en laissant négligemment de nombreuses traces de ses communications sur Internet (inutile de préciser à ce propos que le narcissisme de bien des blogueurs est un trésor pour les exploitants du data *mining* et/ou du s*ocial network analysis*)? ou en consentant de manière irréfléchie à la transmission de ses données personnelles à des entreprises commerciales, en contrepartie de l'accès aux prestations des opérateurs de réseaux sociaux[13]?

3.2. *Le détournement de finalité*

Plus grave, le *Big data* traite souvent des données recyclées. Collectées à l'origine pour un but précis et validé par la personne concernée, elles se voient intempestivement reliées et corrélées avec des informations de provenances différentes; et ce pour d'autres objectifs, inconnus de la personne concernée.

Cette imprévisibité de la finalité du traitement est le lot des mégadonnées. Comme l'a relevé un informaticien américain: «The challenge of analysing *Big data* is coping with abundance, exhaustivity and variety, timeliness and dynamism, messiness and uncertainty, high relationality, and the fact that much of what is generated has no specific question in mind or is a by-product of another activity»[14].

3.3. *Un droit d'accès affaibli*

La faculté de traiter en masse des données non structurées, d'origine et de nature diverses, non seulement occulte les circonstances entourant le traitement mais aussi rend le droit d'accès illusoire. Il était aisé de renseigner la personne concernée, à sa demande, sur les données personnelles traitées, sur leur origine et sur l'objectif du traitement lorsque celles-ci étaient réunies dans des fichiers ordonnés et persistants. Il l'est beaucoup moins lorsque les agrégats de données constitués pour l'analyse sont hétéroclites et volatiles.

3.4. *L'anonymisation: une protection trompeuse?*

Le champ d'application des législations sur la protection des données repose sur la distinction entre données personnelles et données non personnelles. Les premières tombent sous le coup du régime juridique (très contraignant) de la protection des données; les secondes y échappent.

Ressortissent à la catégorie des données personnelles toutes les informations concernant une personne physique identifiée ou identifiable, ce qui inclut les informations qui permettent, par simple croisement ou déduction, de remonter à une personne déterminée (p. ex. l'adresse IP d'un ordinateur, laquelle donne la possibilité de découvrir le propriétaire de la machine)[15]. Cela dit, une personne physique «n'est pas considérée comme identifiable si cette identification nécessite des délais, des coûts et des activités déraisonnables»[16]. En conséquence, il est possible d'échapper au régime de la protection des données en anonymisant ou en pseudonymisant (cryptage des identifiants) les données personnelles, pour autant qu'il ne soit pas aisé de relier les données aux personnes concernées.

A l'heure du *Big data*, la distinction, pourtant fondamentale, entre données personnelles et donnés non personnelles, perd de son sens. Les

puissants algorithmes de traitement des mégadonnées permettent en effet une aisée réidentification, par la rapide corrélation de gigantesques volumes de données non personnelles. Au point que nombre d'informaticiens soutiennent qu'il n'y a plus de données non personnelles, car toute donnée est désormais potentiellement personnelle: «Much of this data seems impersonal. But it's not. What modern data science is finding is that nearly any type of data can be used, much like a fingerprint, to identify the person who created it: your choice of movies on Netflix, the location signals emitted by your cell phone, even your pattern of walking as recorded by a surveillance camera. In effect, the more data there is, the less any of it can be said to be private, since the richness of that data makes pinpointing people «algorithmically possible!»[17].

En conséquence, l'anonymisation, voire la pseudonymisation deviennent des mesures de sauvegarde de la vie privée trompeuses, car elles peuvent être aujourd'hui facilement réduites à néant par le traitement massif de données[18]: plus le volume de données traitées est grand, plus nombreuses sont les corrélations permettant la réidentification[19].

4. Réformer le cadre juridique

4.1. *Un cadre dépassé*

Il est indéniable que l'opacité du *Big data* comme le recyclage abusif des données portent atteinte aux fondements de la protection des données. Sont au premier chef mis-à-mal, les principes cardinaux de transparence et de finalité du traitement, ainsi que celui de minimisation des données. Ces atteintes sont graves, car elle génère une profonde perte de maîtrise de la personne concernée sur ses données, perte de maîtrise aggravée par l'affaiblissement du droit d'accès. N'étant plus en mesure de savoir ce qui se trame, la personne concernée est réduite à l'état d'objet de monitorage et de profilage à tous crins: elle ne peut exercer ni son droit de s'opposer à des traitements qu'elle juge inopportun, ni celui de demander la rectification des données erronées ou la destruction des données excessives.

On sait que le cadre juridique international de la protection des données est en passe d'être modernisé (cf. supra 2). Si tout le monde s'accorde à dire qu'une révision est nécessaire, un profond clivage se fait jour entre ceux qui estiment que les nouveaux textes répondent aux attentes – même si ils ne prennent pas directement en compte les mégadonnées (ni le préambule au règlement général (UE) 2016/679 ni le rapport explicatif au protocole d'amendement de la convention 108 du Conseil de l'Europe ne les mentionnent…) – et ceux qui sont d'avis qu'il faudrait en faire plus.

4.2. *La modernisation en cours*

L'objectif premier de la modernisation est de redonner à la personne concernée le pouvoir de contrôler l'utilisation qui est faite de ses données. Lors du lancement du processus de révision de la directive 95/46, le 25 janvier 2012, Mme Viviane Reding, commissaire européenne chargée de la justice, a été on ne peut plus clair sur ses intentions : «La protection des données à caractère personnel est un droit fondamental reconnu à tous nos concitoyens, mais ceux-ci n'ont pas toujours le sentiment de maîtriser entièrement les données à caractère personnel les concernant. Nos propositions législatives contribueront, dès lors, à susciter la confiance dans les services en ligne parce que les utilisateurs seront mieux informés de leurs droits et auront une plus grande maîtrise des informations qui les concernent».

Dans cette perspective, le législateur européen compte sur le succès de trois mesures spécifiques qu'il a introduites dans le règlement général:
– le renforcement de la qualité du consentement. D'abord l'individu doit bénéficier d'une information claire et intelligible sur les données traitées et la finalité du traitement. Ensuite, son consentement au traitement doit être dénué de toute ambiguïté[20]; à cet égard le préambule du règlement général, à son chiffre 32, souligne que: «Il ne saurait dès lors y avoir de consentement en cas de silence, de cases cochées par défaut ou d'inactivité. Le consentement donné devrait valoir pour toutes les activités de traitement ayant la ou les mêmes finalités. Lorsque le traitement a plusieurs finalités, le consentement devrait être donné pour l'ensemble d'entre elles»;
– l'institution d'un droit à l'oubli qui permet à la personne concernée d'exiger l'effacement de donnés qui ne sont plus nécessaires au regard des finalités pour lesquelles elles ont été collectées[21];
– l'obligation de tenir compte de la protection des données dès la conception d'un produit ou d'un service (*privacy by design*) et par défaut (*privacy par default*). Concrètement: les responsables de traitement devront être conscients des atteintes à la vie privée que sont susceptibles de porter leurs produits ou services et configurer l'architecture du système de telle manière que l'impact négatif soit le plus petit possible; en particulier, ils devront veiller à limiter la quantité de données traitée dès le départ pour respecter le principe de «minimisation «des données[22]. Par exemple, la configuration par défaut des réseaux sociaux devrait être telle que les messages ou les images ne soient partagés qu'avec un cercle restreint et choisi d'individus et non avec l'ensemble des internautes[23].

A ces trois mesures (que consacre également le protocole de modernisation de la convention 108 du Conseil de l'Europe[24]), vient s'ajouter une extension du champ d'application du régime juridique de la protection des données: le respect du règlement général s'impose non seulement aux entreprises situées sur le territoire de l'Union européenne, mais aussi à celles qui, quoique située à l'extérieur de ce périmètre,

conduisent des activités susceptibles d'affecter des résidents européens. Répondent à ce dernier critère: l'offre de biens ou de services à des consommateurs dans l'UE ou le profilage de personnes dans l'UE[25]. Cette extension est décisive si l'on veut enfin discipliner les grands opérateurs américains (on songe en particulier à Google, Facebook et Yahoo) qui jusqu'alors ont toujours argué de leur localisation extra-européenne pour échapper aux exigences de la protection de données[26].

Pour la plupart des experts, les principes fondamentaux de la protection des données, renforcés par quelques aménagements ponctuels, à l'instar de mesures qui viennent d'être présentées, devraient suffire à brider les mégadonnées. A témoin cette déclaration qui émane du Groupe de travail 29, qui réunit les dirigeants des autorités nationales de protection des données des Etats membres de l'Union Européenne: «The Working Party acknowledges that the challenges of *Big data* might require innovative thinking (…). However, at this stage, it has no reason to believe that the EU data protection principles, as they are currently enshrined in Directive 95/46/EC, are no longer valid and appropriate for the development of *Big data*, subject to further improvements to make them more effective in practice»[27].

5. Une nouvelle approche

Nous sommes d'avis que la modernisation en cours sera impuissante à appréhender les conséquences du changement radical dans la manière de traiter les données induite par le traitement des mégadonnées. Des mesures plus drastiques s'imposent; plus drastiques, car elles supposent d'une part une réorientation complète des objectifs de la protection des données, d'autre part une refonte des procédures de création du droit pertinent.

5.1. *Réorienter l'objectif de la protection des données*

La protection des données est conçue comme un instrument destiné à garantir à la personne concernée la pleine maîtrise sur ses données[28]. Une conception volontariste qui trouve son origine dans un jugement de la Cour suprême allemande, lequel consacra, en 1983, l'existence d'un droit fondamental à «l'autonomie informationnelle»[29]. En bref, la personne concernée est investie du droit de décider souverainement de l'utilisation faite par des tiers de ses données personnelles et, le cas échéant, de s'opposer à un traitement qui ne lui convient pas. En d'autres termes, il appartient à la personne concernée de déterminer quelles données à son sujet peuvent être traitées, par qui et dans quel but. Sans aller jusqu'à reconnaître formellement un droit à l'autonomie informationnelle, la Cour de justice de l'Union Européenne et la Cour eu-

ropéenne des droits de l'Homme se montrent, elles-aussi, de plus en plus soucieuses d'assurer à la personne concernée la possibilité de contrôler réellement l'usage qui est fait de ses données.

Contrôle, maîtrise, autonomie sont des belles paroles qui se révèlent aujourd'hui décalées de la réalité. Avec le *Big data*, qui, on l'a vu (cf. supra 3), se joue des principes fondateurs de la protection des données, on assiste à une véritable «érosion du droit à l'autodétermination en matière de données personnelles»[30]. Si l'on veut vraiment protéger la vie privée, il faut repartir sur de nouvelles bases; l'approche «empowerment»de la personne concernée doit être abandonnée au profit d'une approche fondée sur les risques posés par le traitement des mégadonnées. En d'autres termes, l'attention du législateur doit porter moins sur ce que veut la personne concernée que sur ce que font les responsables des traitements. C'est à eux que le législateur doit s'adresser en leur prescrivant des comportements propres à prévenir les abus dommageables à la vie privée.

5.2. *Une réforme institutionnelle*

Des prescriptions sont donc attendues. Pour être efficaces, elles devront être adéquates, circonstanciées et à jour. Trois exigences qualitatives qui sont un réel challenge pour le législateur en prise avec le monde instable, imprévisible et évolutif des nouvelles technologies d'information et de communication.

Pour atteindre le triple objectif de pertinence, d'actualité et de précision normatives dans un domaine aussi incertain que celui du *Big data*, il n'y a qu'une seule solution: décharger le législateur légitime classique (lent et dépassé par un progrès scientifique dont les tenants et aboutissants lui échappent trop souvent) et déléguer la tâche de produire les normes protectrices de la vie privée à une autorité spécifique, à créer au niveau européen. Composée d'experts et de représentant des diverses parties prenantes (notamment des opérateurs de plateformes Internet, de la société civile et des usagers), cette autorité devrait disposer des connaissances de pointe nécessaires afin de réguler efficacement le traitement des données ainsi que l'agilité susceptible d'adapter rapidement les prescriptions aux incessantes innovations technologiques.

6. Le mot de la fin

La présente contribution était dédiée aux menaces que fait planer le *Big data* sur la vie privée. On ne saurait conclure sans relever que celui-ci peut porter des atteintes tout aussi graves à d'autres droits fondamentaux comme la dignité humaine ou l'égalité de traitement.

Ainsi l'analyse statistique des mégadonnées a ouvert la voie à des prédictions portant sur le comportement futur des individus (predictive analytics). Ces prédictions, qui bénéficient d'une apparence trompeuse de fiabilité et d'incontestabilité, peuvent déboucher sur des décisions négatives concernant des personnes (refus d'emploi, fermeture d'un compte bancaire ou rejet de requêtes d'aides financières par exemple). Des décisions négatives d'autant plus regrettables qu'elles sont souvent automatisées. L'absence de possibilité pour la personne concernée de faire valoir son point de vue dans la procédure décisionnelle peut alors être à l'origine de traitements discriminatoires ou d'exclusions arbitraires. Faute de place, on se contentera de simplement relever que le règlement général contient des dispositions tendant à restreindre les cas de prises de décision automatique, fondées sur le profilage[31]. C'est déjà cela, mais d'autres mesures seront inévitables, si l'on veut éviter qu'avec les mégadonnées les statistiques ne prennent le pouvoir.

Notes

[1] Directive 95/46/CE du 24 octobre 1995, relative à la protection des personnes physiques à l'égard du traitement des données à caractère personnel et à la libre circulation de ces données.

[2] Règlement (UE) 2016/679 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 (ci-après, le règlement général).

3 Passer d'une directive à un règlement est la marque d'une volonté ferme de sauvegarder efficacement la vie privée en garantissant l'uniformité des standards et des modalités de protection de données sur tout le territoire de l'Union Européenne.

[4] La croissante numérisation des diverses interactions sociales des individus a conduit à une prolifération vertigineuse du nombre de données produites et partant des possibilités de traçabilité des individus (comme le dénonçait ironiquement l'informaticien américain Bruce Schneier sur son blog, en 2009 déjà: «Welcome to the future, where everything about you is saved. A future where your actions are recorded, your movements are tracked, and your conversations are no longer ephemeral. A future brought to you not by some 1984-like dystopia, but by the natural tendencies of computers to produce data»). Une comparaison suffira pour donner un ordre de grandeur des volumes gigantesques de données en jeu: depuis 2010 il a été produit tous les deux jours autant de données que du début de l'humanité jusqu'en 2003.

[5] Ainsi, par exemple, le Groupe de travail 29, qui réunit les dirigeants des autorités nationales de protection des données des Etats membres de l'Union Européenne, a salué les aspects bénéfiques des mégadonnées: «Many individual and collective benefits are expected from the development of *Big data* (…). The Working Party would naturally support genuine efforts at EU or national levels which aim to make these benefits real for individuals in the EU, whether individually or collectively» (Statement of the WP29 on the impact of the development of *Big data* on the protection of individuals with regard to the processing of their personal data in the EU, 16 septembre 2014, p. 2).

[6] On l'aura constaté, la présente contribution, par simplification, met protection de la vie privée et protection des données sur le même plan. Nous sommes toutefois conscients que les deux domaines ne sont pas exactement superposables; la protection des données comporte notamment une dimension de contrôle de l'individu sur ses données personnelles qui ne ressortit pas à la protection de la vie privée stricto sensu (voir sur ce point, Orla Lynskey, «Deconstructing data protection: the 'added-value' of a right to data protection in the EU legal order», *International and Comparative Law Quarterly*, 2014, p. 569 ss).

[7] La protection des données est avant tout une affaire européenne. C'est sur ce continent que sont nées les premières lois en la matière et qu'elles sont, aujourd'hui encore, les plus strictes. Sur le clivage entre une Europe au niveau de protection très élevé et le reste du monde, plus laxiste, voir Bertil Cottier, «Gouvernance d'Internet: Protection de la vie privée et des données personnelles», *Revue suisse de droit international et européen*, 2016, p. 258 ss.

[8] Cette disposition «constitutionnelle» non seulement consacre expressément un droit à la protection des données personnelles, mais encore en défini les contours généraux.

[9] Convention 108 du 28 janvier 1981 pour la protection des personnes à l'égard du traitement automatisé des données à caractère personne (ci-après convention 108). Tous les pays membres du Conseil de l'Europe ont ratifié ce texte.

[10] Voir notamment l'art 5 al. 1 du règlement général.

[11] Voir l'art. 5 al. 1 de la version consolidée de la convention 108.

[12] Par métadonnées, on entend des renseignements accessoires sur le volume et la durée d'une communication, sur la localisation d'une personne ou encore sur la provenance d'une information.

[13] Le consentement à la transmission des données personnelles résulte souvent d'une clause reléguée au fin fond des conditions générales du service convoité. Dans l'immense majorité des cas, il est donné sans avoir même pris connaissance de la teneur des conditions générales applicables, par un simple click de souris sur la rubrique *j'accepte*. Cela dit, les conditions générales des grands opérateurs de plateformes Internet présentent tous les aspects de contrats d'adhésion non négociables. Dès lors, le consommateur n'a guère de choix: c'est à prendre ou à laisser.

[14] Rob Kitchin, «*Big data*, New Epistemologies and Paradigm Shifts», *Big data and Society*, April-June 2014, p. 1ss. Voir aussi Florent Thouvenin, «Erkannbarkeit und Zweckbindung, Grundprinzipien des Datunschutzrechts auf dem Prüfstand von *Big data*», in Rolf H. Weber et Florent Thouvenin (éd.), *Big data und Datenschutz – Gegenseitige Herausforderungen*, Zurich, 2014, p. 61 ss.

[15] Voir notamment l'art. 4 (1) a du règlement général qui définit les données à caractère personnel comme «toute information concernant une personne physique identifiée ou identifiable (personne concernée); est réputée identifiable une personne qui peut être identifiée, directement ou indirectement, notamment par référence à un numéro d'identification ou à un ou plusieurs éléments spécifiques, propres à son identité physique, physiologique, psychique, économique, culturelle ou sociale».

[16] Conseil de l'Europe, *Manuel de droit européen de la protection des données*, Strasbourg 2014, p. 44.

[17] Arvind Narayanan, professeur à l'Université de Princeton, cité par Patrick Tucker, «Has *Big data* Made Anonymity Impossible?», *MIT Technology Review*, 7 mai 2013. Voir aussi NeoLex Avocats, «*Big data*: vers une perte certaine de notre anonymat?» contribution sur le blog *Droit et innovation* (6 mai 2013) et Antoinette Rouvroy, *Des données et des hommes. Droits et libertés fondamentaux dans un monde de données massives*, Conseil de l'Europe, Strasbourg 2016, p. 24 ss.

[18] À terme, le *Big data* pourrait également vider de toute signification les lois sur la protection des données elles-même. Comme le fait remarquer Bruno Baeriswyl (*«Big data» ohne Datenschutz-Leitplanken*, digma 2013, p. 16), le fait que toutes les données renferment désormais un potentiel d'identification les soumet toutes, quelles qu'elles soient, au régime de la protection des données, une conséquence aussi absurde que gravement attentatoire à la liberté de l'information.

[19] Comme le souligne Antoinette Rouvroy (op. cit., note 18, p. 25) : «La valeur de chaque donnée n'est plus, dans le contexte des *Big data*, contenue en elle-même, mais est essentiellement de nature relationnelle: ce sont les (cor)relations entre données qui leur confèrent une utilité, une valeur, et aussi, éventuellement, un caractère plus ou moins sensible».

[20] Art. 7, 13 et 14 règlement général.

[21] Art. 17 règlement général.

[22] Art. 25 règlement général.

[23] Projet de rapport explicatif à au protocole d'amendement de la convention 108 du Conseil de l'Europe, chiffre 86 (août 2016).

[24] Voir notamment ses articles 5 al.2 (consentement éclairé), 7bis (transparence du traitement) et 8bis (*privacy by design and by default*) de la version consolidée de la convention 108 (septembre 2016).

[25] Art. 3 al. 2 règlement général.

[26] Avec il est vrai de moins en moins de succès depuis l'arrêt Google de la Cour de justice de l'Union européenne C-131/12 (14 mai 2014) qui a contraint Google Inc. (Californie) à déréférencer des articles de presse concernant un citoyen espagnol.

[27] Op. cit. (note 6) p. 2.

[28] Antoinette Rouvroy et Yves Poulet, «The Right to Informational Self-Determination and the Value of Self Development», in: Serge Gurwith et alii (éd.), *Reinventing Data Protection*, Berlin 2009, p. 69ss.

[29] Arrêt du 15 décembre 1983 (Az. 1 BvR 209/83, *Volkszählungsurteil*).

[30] Alexandre Flückiger, «L'autodétermination en matière de données personnelles : un droit (plus si) fondamental à l'ère digitale ou un nouveau droit de propriété?», *Pratique juridique actuelle*, 2013, p. 837s.

[31] Voir l'art. 22 Règlement. Pour plus d'information sur les analyses prédictives, voir Antoinette Rouvroy, op. cit. (note 18), p. 36ss.

Summary

Big data and privacy protection
by Bertil Cottier

Big data has revolutionized the ways information is collected, combined and evaluated. Relying on huge amounts of data originating from a multiplicity of sources, big data analytics allow for the unprecedented monitoring of citizens. Core data-protection principles of transparency, proportionality and purpose limitation are put into question. European legislators have reacted by strengthening relevant data protection instruments. In particular, the General Data Protection Regulation (EU) 2016/679, which will come into force in 2018, enables individuals to better control the use of their data. Though innovative, like the «privacy by design» obligation or the right to be forgotten, the improvements provided for are not sufficient to tackle formidable scientific progress in data aggregation. Hence, the big data revolution calls for a substantive change in the way data protection is conceived and regulated: a risk-based approach should supplant the traditional but deceptive informational autonomy approach.

Keywords: privacy, data protection, regulation, monitoring, informational autonomy

JEL code: C80, D18, K20, K36