

## Le Big Data est-il compatible avec le RGPD ?

Le Big Data est le nouvel outil chéri des directions marketing, surtout lorsqu'il est couplé avec ce qu'il est convenu d'appeler l'intelligence artificielle. Effectivement, ces techniques sont prometteuses : meilleures décisions, moindre risque, diminution des coûts... Mais quel est leur impact sur la protection des données personnelles ? Dans quelle mesure ces méthodes nouvelles sont-elles utilisables en Europe, où le respect du RGPD est une obligation légale ?

Pour répondre à ces questions, nous allons tout d'abord regarder dans les grandes lignes comment fonctionne le Big Data ; puis nous indiquerons les éléments juridiques qui peuvent rendre incompatibles Big Data et RGPD ; enfin nous donnerons quelques pistes permettant d'éviter ces chausse-trappes.

Le Big Data, c'est l'analyse de grands ensembles de données afin d'en extraire des connaissances nouvelles. À la base, c'est de la statistique ; on veut, à partir des données disponibles – plus il y en a, mieux ça marche - trouver des tendances, classifier des comportements, bref extraire des informations cachées. La technique la plus récente dans ce domaine est connue sous le nom d'"apprentissage profond" (que les consultants préfèrent bien sûr appeler "*deep learning*", même si un de ses inventeurs est français...)

Alors, nous l'avons dit, il faut beaucoup de données. Les entreprises (et sûrement les États) se mettent donc à accumuler les données. Sans trop savoir ce qu'ils vont en faire, au moment où ils les collectent. Et c'est là que le bât blesse : lorsqu'il s'agit (comme c'est souvent le cas) de données personnelles, le RGPD impose (article 5.1b) une limitation des finalités, qui passe par une description des usages prévus lors de la collecte. Et en plus, il faut minimiser les données recueillies (article 5.1c) ! Difficile apparemment de concilier ces exigences avec le désir légitime de multiplier les analyses de données sans savoir initialement ce qu'on cherche... et qui en outre va certainement évoluer au fil des différentes étapes !

Heureusement, le législateur a prévu dans ce même article des marges de flexibilité, notamment lorsqu'il s'agit d'une exploitation statistique des données – or nous savons que c'est la base de nombreux traitements qui exploitent de grandes quantités de données.

L'article 89 du RGPD précise les conditions nécessaires pour pouvoir bénéficier de ces marges de manœuvre. En substance, il s'agit de s'assurer que les droits et libertés de la personne concernée (dans le jargon du RGPD, c'est celle dont on traite les données personnelles) sont respectés. On peut pour cela utiliser en particulier des techniques d'anonymisation ou de pseudonymisation.

## Le Big Data est-il compatible avec le RGPD ?

Ce bref article ne prétend pas faire le tour de la question, et il reste de nombreuses sources d'incertitude juridique qu'il faut circonscrire. Nous serons heureux d'en discuter avec vous !

**Lúcia DA CUNHA et Jackie POSTMA**



\* Les informations contenues dans cet article n'engagent que ses auteurs. Le rôle du CONJUR se limite à la divulgation des productions intellectuelles de ses membres, n'exerçant aucun contrôle sur le fond du sujet.