

big data au passé, présent et futur

On sait tout, sur les réseaux sociaux, de votre passé. On prétend prédire ce qu'il va vous arriver grâce à des programmes prédictifs. La loi pour une République numérique et le règlement général sur la protection des données personnelles tentent aujourd'hui de dresser des paravents pour protéger votre vie privée.

Un mauvais algorithme avec beaucoup de data serait supérieur à un bon algorithme avec peu de data estiment des spécialistes de l'intelligence artificielle. C'est dire la puissance du big data qui permet de tout savoir et prétend tout prévoir.

Google enregistre toutes nos requêtes, Facebook sait ce que l'on aime, Amazon connaît notre localisation, iTunes répond à nos goûts musicaux et YouTube les devine. Le plus terrible est que toutes ces informations détenues par les Gafam (1) sont conservées depuis notre premier clic sur leur site et qu'ils entendent les garder ad vitam aeternam.

1. plongée dans le passé

Le 18 décembre 2013 à 9 heures 24, Judith Duportail écrit son premier « Hello » sur Tinder. En mars 2017, elle obtient de ce réseau social de

rencontre les archives de la totalité de ses connexions et propos échangés lors de 870 « matches ». Tinder lui a envoyé, à sa demande insistante, l'équivalent de 800 pages dans lesquelles toute une partie de sa vie intime remonte à la surface, extirpée de ses souvenirs enfuis. Tous ses « like » sur Facebook, tous ses liens sur des photos d'Instagram, tous les contenus des 1 700 messages adressés via Tinder à des dizaines d'hommes et où elle retrouve tous ses copier-coller. C'est Judith Duportail elle-même qui nous révèle l'abîme de son passé dans le journal The Guardian où elle est journaliste (2). Elle a eu beaucoup de peine à obtenir la photocopie de son passé, car les réseaux sociaux ont du mal à comprendre qu'ils ne sont pas propriétaires de nos informations personnelles. Pour pouvoir exercer son droit d'accès prévu par le droit européen (3), elle a dû recourir à l'expertise d'un activiste de la protection de la vie privée et à un avocat spécialiste des droits de l'Homme !

Puisque les serveurs savent tout de nous hier, certains en ont conclu que l'on peut désormais tout connaître de notre demain.

2. incursion dans le futur

En fait, il s'agit d'un futur qui a déjà commencé et qui s'appelle la justice prédictive. La plupart des avocats évaluent aujourd'hui pour leur client les chances de succès avec des marges d'incertitude importantes, faute de statistiques. On n'est jamais sûr de gagner (ni de perdre !). On ne sait pas combien précisément on va obtenir de dommages et intérêts (4).

Des laboratoires de R&D des « law firms » américaines jusqu'aux start-up françaises, nombreux sont ceux qui développent des logiciels d'aide à la décision pour les avocats :

« la loi pour une république numérique prévoit que la mise à disposition des décisions de justice devra respecter la vie privée des personnes concernées »

cartographie des juridictions les plus favorables en fonction de l'objet du litige ; identification et hiérarchisation des arguments juridiques qui sont les mieux reçus par les juges de tel ou tel tribunal ; estimation des indemnités à espérer compte tenu des sommes obtenues dans des contentieux similaires antérieurs. Le magistrat Antoine Garapon estime que les algorithmes « révèlent l'inconscient des institutions juridiques » (5).

la prédiction renforce la prédiction

Les assureurs ne sont pas en reste. Ceux qui fournissent des services de protection juridique utilisent des outils de calcul de probabilité de succès qui amènent 75 % de leurs clients à accepter une solution amiable. Résultat impressionnant lorsque l'on connaît la méfiance des Français envers la conciliation. Ils



ont toujours peur du risque de « *se faire avoir* ». En revanche, si la machine leur prédit le futur, ce risque disparaît et ils acceptent ce futur prédit. En quelque sorte, la prédiction renforce la prédiction. L'assuré accorde donc une importance particulière à la statistique, là où un argumentaire fondé sur quelques jurisprudences éparses échoue bien souvent à le convaincre de renoncer à une action contentieuse (6).

Pour que l'outil de justice prédictive fonctionne, il faut l'alimenter avec beaucoup de données contentieuses. Pour répondre à cette fringale de décisions de justice, la loi pour une République numérique de 2016 a inséré une obligation d'open data dans l'article L.10 du Code de la justice administrative et dans l'article L.111-13 du Code de l'organisation judiciaire. Les décisions des juridictions administratives et judiciaires vont ainsi être mises à la disposition du public à titre gratuit (7). Les décrets d'application sont en attente de publication pour 2018.

menace à la vie privée des justiciables

Face à cette vague d'open data de données brutes, les éditeurs de bases de données privées vont proposer des données enrichies grâce à leur savoir-faire et leur fonds documentaire (par exemple, Lexbase est la plus grande base de jurisprudence française avec à ce jour 4 millions de décisions décomptées et s'y ajoutent des milliers d'articles de doctrine).

Dans un monde où l'information circule en quelques secondes autour de la planète médiatique via les réseaux sociaux, l'open data pour les décisions de justice menace la vie privée des justiciables. Pour prédire qui aura la garde du chien après le divorce, l'algorithme a-t-il besoin de connaître les noms des époux ? Le concept même du big data amène à une réponse positive. Le principe de protection de la vie privée impose une réponse négative. Et c'est bien ce qu'exigent la loi Lemaire du 7 octobre 2016 et le RGPD, règlement général sur la protection des données (personnelles).

3. nos droits sur nos données

La loi pour une République numérique prévoit que la mise à disposition des décisions de justice devra respecter la vie privée des personnes concernées et que cette mise à disposition du public doit être précédée d'une analyse du risque de ré-identification des personnes. La loi demande donc une réelle dépersonnalisation dans les bases de données publiques.

Au-delà du cas particulier de la justice prédictive, et pour revenir à la journaliste Judith Duportail, la présidente de la

« toute personne dispose du droit de décider et de contrôler les usages qui sont faits des données à caractère personnel la concernant, dans les conditions fixées par la présente loi »

Cnil considère que la loi pour une République numérique et le RGPD donnent aux individus la possibilité d'acquérir une souveraineté numérique : une souveraineté individuelle, car les individus gagnent de la maîtrise sur les informations qui les concernent, et une souveraineté territoriale dans la mesure où il remet les acteurs européens à égalité de concurrence avec les Gafam qui seront soumis au droit européen dès lors qu'ils offrent un produit ou un service à un citoyen européen (8).

droit à « l'autodétermination informationnelle »

La loi Lemaire constitue donc incontestablement une avancée pour la protection des droits des personnes. Elle complète aussi l'article 1^{er} de la loi

Informatique et Libertés de 1978 en prévoyant que « *Toute personne dispose du droit de décider et de contrôler les usages qui sont faits des données à caractère personnel la concernant, dans les conditions fixées par la présente loi* ». Elle réaffirme qu'à l'ère numérique, l'individu est le centre de gravité de la législation sur la protection des données en reconnaissant aux personnes en quelque sorte un droit à « *l'autodétermination informationnelle* », y compris la possibilité de prévoir le sort de ses données après sa mort. En outre, un droit à l'oubli pour les mineurs est consacré et les modalités d'information des gens sur ce qui va être fait des données qui les concernent sont renforcées. ■

Sylvain Martin

[Avocat à la Cour. Membre du conseil d'administration de l'Adij]

→ www.sylvain-martin.com

(1) Google, Amazon, Facebook, Apple, Microsoft.

(2) → www.theguardian.com/technology/2017/sep/26/tinder-personal-data-dating-app-messages-hacked-sold.

(3) Directive 95/46 du 24 octobre 1995 remplacée le 25 mai 2018 par le Règlement général sur la protection des données (personnelles) – RGPD – 2016/679 du 27 avril 2016.

(4) À cet égard, l'ordonnance Macron qui plafonne les dommages et intérêts suite à un licenciement abusif coupe l'herbe sous les pieds des logiciels prédictifs qui, sur ce point, ne vont être que des logiciels énonciatifs de la règle.

(5) France Culture le 11 février 2017, « *Justice prédictive, les big data menacent-elles la justice ?* »

(6) Étude réalisée par Predictice, sur un panel d'utilisateurs assureurs de protection juridique. Cette société a conclu un partenariat avec le barreau de Lille début 2017 :

→ www.actualitesdudroit.fr/browse/tech-droit/start-up/9576/analyse-algorithmique-des-donnees-judiciaires-de-l-aide-a-la-decision-a-la-justice-predictive

(7) Loi 2016-1321 du 7 octobre 2016, art. 20 et 21.

(8) Rapport d'activité 2016, page 11.